Unveiling a 36 Billion Solar Mass Black Hole at the Centre of the Cosmic Horseshoe Gravitational Lens

Carlos R. Melo-Carneiro,[1,2]⋆ Thomas E. Collett,[2] Lindsay J. Oldham,[2] Wolfgang Enzi,[2] Cristina Furlanetto[1] and Ana L. Chies-Santos[1]

[1]*Instituto de Física, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, Porto Alegre-RS, 90040-060, Brazil*
[2]*Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Rd, Portsmouth, PO1 3FX, UK*

ABSTRACT

Supermassive black holes (SMBHs) are found at the centre of every massive galaxy, with their masses tightly connected to their host galaxies through a co-evolution over cosmic time. For massive ellipticals, the SMBH mass (M_{BH}) strongly correlates with the central stellar velocity dispersion (σ_e), via the $M_{BH} - \sigma_e$ relation. However, SMBH mass measurements have traditionally relied on central stellar dynamics in nearby galaxies ($z < 0.1$), limiting our ability to explore the SMBHs across cosmic time. In this work, we present a self-consistent analysis combining 2D stellar dynamics and lens modelling of the Cosmic Horseshoe gravitational lens system ($z = 0.44$), one of the most massive galaxies ever observed. Using integral-field spectroscopic data from MUSE and high-resolution imaging from HST, we model the radial arc and stellar kinematics, constraining the galaxy's central mass distribution and SMBH mass. Bayesian model comparison yields a 5σ detection of an ultramassive black hole (UMBH) with $\log_{10}(M_{BH}/M_\odot) = 10.56^{+0.07}_{-0.08} \pm (0.12)^{sys}$, consistent across various systematic tests. Our findings place the Cosmic Horseshoe ~1.5σ above the $M_{BH} - \sigma_e$ relation, supporting an emerging trend observed in BGCs and other massive galaxies. This suggests a steeper $M_{BH} - \sigma_e$ relationship at the highest masses, potentially driven by a different co-evolution of SMBHs and their host galaxies. Future surveys will uncover more radial arcs, enabling the detection of SMBHs over a broader redshift and mass range. These discoveries will further refine our understanding of the $M_{BH} - \sigma_e$ relation and its evolution across cosmic time.

Key words: gravitational lensing: strong – galaxies: kinematics and dynamic – galaxies: evolution – quasars: supermassive black holes

1 INTRODUCTION

Most massive galaxies are believed to host a supermassive black hole (SMBH) at their centre. More importantly, host galaxies and their SMBHs exhibit clear scaling relations, pointing to a co-evolution between the galaxy and the black hole (BH; Kormendy & Ho 2013). The BH mass (M_{BH}) has been shown to correlate with various galaxy properties, such as the bulge luminosity (e.g., Magorrian et al. 1998; Marconi & Hunt 2003; Gültekin et al. 2009), stellar bulge mass (e.g., Laor 2001; McLure & Dunlop 2002), dark matter (DM) halo mass (e.g., Marasco et al. 2021; Powell et al. 2022), and stellar velocity dispersion (e.g., Gebhardt et al. 2000; Beifiori et al. 2009). Notably, the $M_{BH} - \sigma_e$ relation, which links BH mass to the effective stellar velocity dispersion of the host (σ_e), remains tight across various morphological types and BH masses (van den Bosch 2016). Nonetheless, when SMBHs accrete mass from their neighbourhoods, they can act as active galactic nuclei (AGNs), injecting energy in the surrounding gas in a form of feedback. This feedback can be either positive, triggering star formation (Ishibashi & Fabian 2012; Silk 2013), or

negative, quenching galaxy growth (e.g., Hopkins et al. 2006; Dubois et al. 2013).

It is expected that the most massive galaxies in the Universe, such as brightest cluster galaxies (BCGs), host the most massive SMBHs. Indeed, so-called ultramassive black holes (UMBHs; $M_{BH} \geq 10^{10} M_\odot$) have been found in such systems (e.g., Hlavacek-Larrondo et al. 2012). Most of these UMBHs have been measured through spatially resolved dynamical modelling of stars and/or gas. For instance, the UMBH in Holm 15A at $z = 0.055$ ($M_{BH} = (4.0 \pm 0.80) \times 10^{10} M_\odot$; Mehrgan et al. 2019) and the UMBH in NGC 4889 at $z = 0.021$ ($M_{BH} = (2.1 \pm 1.6) \times 10^{10} M_\odot$; McConnell et al. 2012) were both determined using stellar dynamical modelling. However, despite the success of this technique in yielding hundreds of BH mass measurements, the requirement for high-quality spatially resolved spectroscopy poses significant challenges for studies at increasing redshift (see, e.g., Kormendy & Ho 2013, Suplemental Material S1).

Nonetheless, the significance of these UMBHs lies in the fact that many of them deviate from the standard linear $M_{BH} - \sigma_e$ relation (e.g., Kormendy & Ho 2013; van den Bosch 2016). This suggests either a distinct evolutionary mechanism governing the growth of the largest galaxies and their BHs (McConnell et al. 2011), leading

to a significantly steeper relation (Bogdán et al. 2018), or a potential decoupling between the SMBH and host galaxy co-evolution. Populating the high-mass end of the $M_{BH} - \sigma_e$ relation, particularly through direct BH measurements, could help resolve this ongoing puzzle.

Recently, Nightingale et al. (2023), by modelling the gravitationally lensed radial image in the Abell 1201 BCG ($z = 0.169$), was able to measure the mass of its dormant SMBH as $M_{BH} = (3.27 \pm 2.12) \times 10^{10} M_\odot$. This provides a complementary approach to other higher-z probes of BH mass, such as reverberation mapping (Blandford & McKee 1982; Bentz & Katz 2015) and AGN spectral fitting (Shen 2013). Unlike these methods, which require active accretion and depend on local Universe calibrations, the lensing technique offers a direct measurement independent of the SMBH's accretion state.

In this paper, we analyse the Cosmic Horseshoe gravitational lens system (Belokurov et al. 2007), where the lens galaxy is one of the most massive strong gravitational lenses known. The lens galaxy is an early-type galaxy (ETG) at redshift $z_l = 0.44$, possibly part of a fossil group, and is notable for lensing one of its sources into a nearly complete Einstein ring (the Horseshoe). Additionally, a second multiply imaged source forms a radial arc near the centre of the lens galaxy.

Thanks to the radial image formed very close to the centre, the inner DM distribution of the Cosmic Horseshoe can be studied in detail, as demonstrated by Schuldt et al. (2019). By simultaneously modelling stellar kinematics from long-slit spectroscopy and the positions of the lensed sources, Schuldt et al. (2019) found that the DM halo is consistent with a Navarro-Frenk-White (NFW; Navarro et al. 1997) profile, with the DM fraction within the Einstein radius (R_{Ein}) estimated to be between 60% and 70%. Moreover, their models include a point mass at the galaxy's centre, reaching values around $\sim 10^{10} M_\odot$, which could represent a SMBH; however, they did not pursue further investigations into this possibility.

Using new integral-field spectroscopic data from the Multi Unit Spectroscopic Explorer (MUSE) and imaging from the Hubble Space Telescope (HST), we conducted a systematic modelling of the Cosmic Horseshoe system to reassess the evidence for an SMBH at the heart of the lens galaxy. We performed a self-consistent analysis of both strong gravitational lensing (SGL) and stellar dynamics, which demonstrated that the presence of an UMBH is necessary to fit both datasets simultaneously. Additionally, we found that including the radial arc improves the accuracy of the BH mass estimation, compared to models based solely on stellar dynamics.

This paper is structured as follows: In Section 2, we present the HST imaging data and MUSE observations, along with the kinematic maps used for the dynamical modelling. Section 3 briefly summarises the lensing and dynamical modelling techniques, including the multiple-lens-plane formalism, the approximations adopted in this work, and the mass profile parametrisation. In Section 4, we present the results from our fiducial model and alternatives models, which we use to address the systematics on the BH mass. In Section 5 we discuss our results and present other astrophysical implications. Finally, we summarise and conclude in Section 6.

Throughout this paper, we adopt the cosmological parameters consistent with Planck Collaboration et al. (2016): $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, $\Omega_{b,0} = 0.0486$, and $H_0 = 67.74$ km s^{-1} Mpc^{-1}.
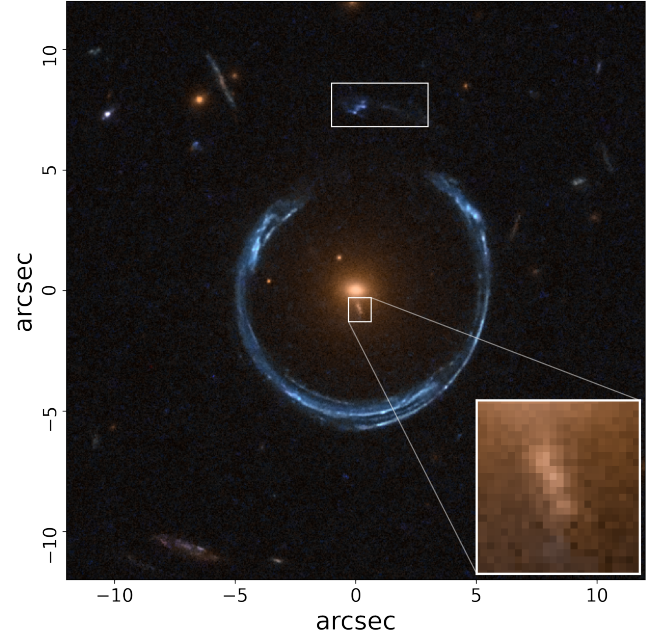
Figure 1. HST/WFC3 colour composite image of the Cosmic Horseshoe, created using the F814W, F606W, and F475W filters. The system is composed by the main deflector ($z_l = 0.44$); the eponymous Einstein ring of the Cosmic Horseshoe ($z_{s2} = 2.381$); and the radial arc and its counter-image ($z_{s1} = 1.961$), both highlighted. The inset shows the radial arc. It is worth noting that, even in the F475W filter where s1 is brighter, the radial arc remains very faint. The figure is oriented such that east is up and north is right.

2 DATA

The Cosmic Horseshoe (SDSS J1148+1930) was first identified by Belokurov et al. (2007) as part of the Sloan Digital Sky Survey (SDSS). The primary lens is a massive ETG at $z_l = 0.44$, with an estimated enclosed mass within R_{Ein} of $\sim 5 \times 10^{12} M_\odot$ (Dye et al. 2008; Schuldt et al. 2019). The radial arc and its counter-image correspond to a source at redshift $z_{s1} = 1.961$ (s1, hereafter), while the tangential arc is a star-forming galaxy (James et al. 2018) at redshift $z_{s2} = 2.381$ (s2, hereafter). In Fig. 1, the radial arc and its counter-image are highlighted within white boxes, with a zoomed-in view of the radial image displayed in the inset.

2.1 HST imaging

The HST images used in this work were obtained with the Wide Field Camera 3 (WFC3) and downloaded from the Hubble Legacy Archive[1] (HLA). Observations with the F475W, F606W, F814W, F110W, and F160W filters were conducted in May 2010[2], while the F275W filter data were taken in November 2011[3]. The data reduction followed the HLA pipeline, which employs the DrizzlePac[4] to process the images. This includes the combination of multiple exposures, correction for geometric distortion, subtraction of the sky background, and removal of cosmic rays. The final science images for the UVIS filters (F275W, F475W, F606W, and F814W) have a pixel

scale of 0.04″, while the IR filters (F110W and F160W) provide images with a pixel scale of 0.13″.

We made use of images in the F475W and F814W filters for our analysis. The F475W band was selected for lens modelling, as the radial arc appears bluer and more distinct from the main-lens in this filter. Conversely, the F814W band was used to trace the light distribution and stellar mass of the primary deflector (see Section 3), as the main-lens is brighter in this filter and the radial arc is not visible, minimising contamination from s1. Both images were aligned using the `Astroalign` software (Beroiz et al. 2020).

To construct the point spread function (PSF) for each filter, we identified non-saturated stars from the Gaia DR2 catalogue (Arenou et al. 2018) and performed interactive PSF modelling using `PSFr` (see, e.g., Birrer et al. 2019). As in Schuldt et al. (2019), the noise map for each pixel i was calculated by combining the background noise σ_{bkgd} and Poisson noise in quadrature: $\sigma^2_{rms,i} = \sigma^2_{bkgd} + \sigma^2_{Poisson,i}$. The background level was estimated as a constant value, measured from an empty region of sky near the main deflector, using the `astropy` sigma-clipping method. The Poisson noise was calculated from the effective exposure map and the intensity counts for each pixel.

2.2 MUSE observations and kinematical map

The integral-field spectroscopic observations were conducted using the VLT/MUSE instrument across three separate visits[5] and retrieved from the ESO Science Archive Facility[6]. The data covers a spectral range of $4650 - 9300$Å, sampled at 1.25Å/px, with a mean spectral resolution of ~ 2.6Å at full width at half-maximum (FWHM; $\sigma \sim 50\,km\,s^{-1}$). The spatial pixel scale is 0.2″, and the seeing during observations was 0.8″. Data reduction followed the ESO Phase 3 Data Release, utilising the MUSE pipeline (Weilbacher et al. 2016). The MUSE data cube was aligned with HST images by generating a collapsed image from the cube and using `Astroalign` to register it with the F475W HST observation.

To extract stellar kinematics from the MUSE data cube, we selected all pixels with a signal-to-noise ratio (SNR) greater than 2.5, excluding regions exhibiting emission from s2. Pixels corresponding to the radial source position were also inspected, but no contribution from s1 was identified in the spectra. The remaining pixels were spatially binned using the Voronoi binning method of Cappellari & Copin (2003) to achieve a minimum SNR of 15. The SNR was calculated as the ratio of the average signal to the average noise in the rest-frame spectral range $5600 - 7600$ Å. Using the continuum rest-frame range $6000 - 6200$ Å, produced negligible differences in the results.

The mean velocity (v) and velocity dispersion (σ_v) in each Voronoi bin were measured using the penalized pixel fitting method, as implemented in the `pPXF` software (Cappellari & Emsellem 2004). For templates, we adopted the X-shooter Spectral Library SSP models (XLS-SSP; Verro et al. 2022), which offer a resolution of ~ 13 $km\,s^{-1}$ and a wavelength coverage from 350 to 2480 nm. The XLS-SSP library was selected due to its high resolution, enabling convolution with the MUSE instrumental resolution after de-shifting the galaxy spectra. Each Voronoi bin was fitted over the wavelength range $5600 - 7600$ Å (galaxy-frame), with emission lines within this range masked. Uncertainties for v and σ_v were determined through Monte

Carlo perturbations of the best-fit model, performing 200 realisations and taking the standard deviation as the uncertainty.

We computed the velocity second moment as $v_{rms} = \sqrt{v^2 + \sigma_v^2}$, finding that rotational velocities are negligible, indicating the galaxy is dominated by the velocity dispersion.

The observed v_{rms} map and radial profile are shown in Fig. 2. The profile is nearly flat in the galaxy's central regions ($< 0.5″$) but increases for $r > 0.75″$. This rising behaviour at larger radii was previously reported by Spiniello et al. (2011) using long-slit spectroscopy of the Cosmic Horseshoe and is a common feature among BCGs (e.g., Newman et al. 2015; Smith et al. 2017b). Notably, the outermost bins exhibit larger error bars, reflecting the low SNR in these regions.

The effective velocity dispersion of the main-lens was determined by co-adding all spectra within the galaxy's effective radius ($R_e = 2.10″$)[7] and fitting the integrated spectrum with `pPXF`, as outlined earlier. This analysis yielded $\sigma_e = 366 \pm 6\,km\,s^{-1}$. We confirmed that using the flux-weighted method present in Emsellem et al. (2007) lead to a similar result.

3 METHODS

We construct a fully self-consistent mass model for the main lensing galaxy by jointly modelling its SGL effect and spatially resolved stellar velocity dispersion. This combined approach has been successfully applied in previous studies to constrain the mass profiles of ETGSs in tests of modified gravity (e.g., Collett et al. 2018; Melo-Carneiro et al. 2023) and to investigate the distribution of baryonic and DM content within galaxies (e.g., Barnabè et al. 2012; Wang et al. 2022). By simultaneously leveraging both methods, we mitigate degeneracies that arise when using them independently, such as the mass-anisotropy degeneracy (e.g., Gerhard 1993) and the mass-sheet degeneracy (e.g., Gorenstein et al. 1988).

3.1 Gravitational Lensing

3.1.1 Formalism

For a single-source plane lens configuration, the source plane position, $\boldsymbol{\beta}$, relates to the observed lensed position, $\boldsymbol{\theta}$, via the lens equation:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\alpha}$ is the reduced deflection angle, given by:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\boldsymbol{\theta}' \kappa(\boldsymbol{\theta}') \frac{(\boldsymbol{\theta} - \boldsymbol{\theta}')}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2}. \tag{2}$$

Here, $\kappa = \Sigma/\Sigma_{crit}$ represents the surface mass density of the lens scaled by the critical surface density, Σ_{crit}, defined as:

$$\Sigma_{crit} = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}}, \tag{3}$$

where D_{ij} is the angular diameter distance between observer, lens, and source.

One can notice that changing the source redshift, only the angular

[7] The effective radius was determined using the galaxy's MGE surface brightness model. See Section 3.3 for more details.
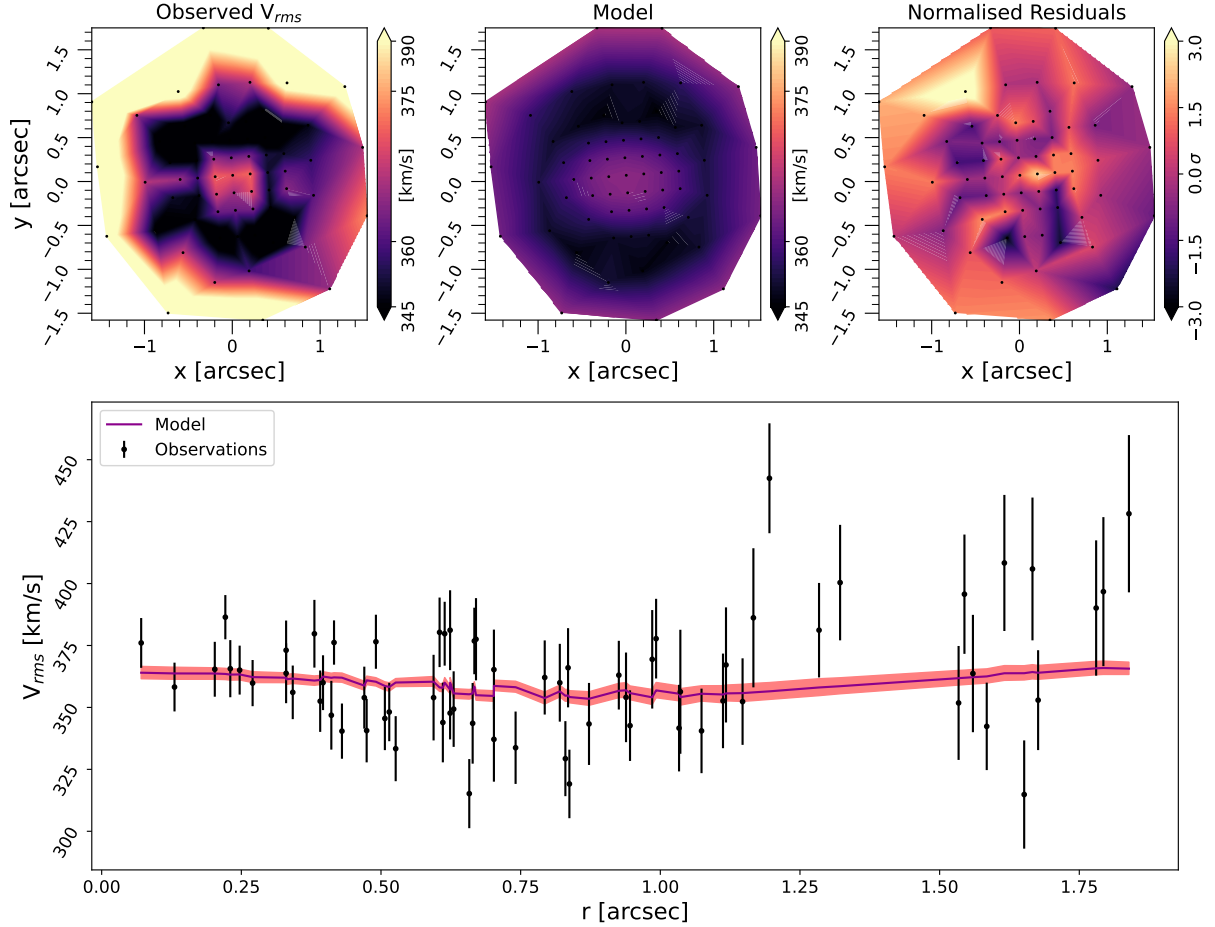
Figure 2. Stellar dynamics fiducial model. The top panels show the observed $v_{\rm rms}$ kinematic map (left), the median kinematic model (centre), and the normalised residuals (right). The bottom panel presents the radial kinematic profile (black dots) alongside the median model and its 1σ credible region. The black dots in the top panels mark the centroids of the Voronoi bins.

diameter distances are changed in the reduced deflection angle, Eq. 2. For two light rays passing through the same point in the lens plane, but originating from different source planes, the relationship between their deflection angles can be expressed as a scaling factor, η, determined by the ratio:

$$\frac{\alpha_1}{\alpha_2} = \frac{D_{ls1}}{D_{s1}}\frac{D_{s2}}{D_{ls2}} \equiv \eta, \quad \text{with } z_{s2} > z_{s1}. \tag{4}$$

In cases where multiple sources are aligned along the line-of-sight (LOS), the gravitational field of the first source lenses the light from the second source before it is further deflected by the main-lens. To account for these effects, a full multiple-lens-plane formalism is required (Schneider et al. 1992). The single-plane lens equation (Eq. 1) can then be generalised into the compound lens equation:

$$\theta_j = \theta_1 - \sum_{i=1}^{j-1} \eta_{ij}\alpha_i(\theta_i), \tag{5}$$

where θ_j is the angular position of a light ray in the $j^{\rm th}$ plane, and α_i is the angular deflection caused by the $i^{\rm th}$ lens acting on rays originating from the furthest redshift source plane. The factor η_{ij} is the scaling factor as defined in Eq. 4, i.e.,

$$\eta_{ij} = \frac{D_{ij}}{D_j}\frac{D_s}{D_{is}}, \tag{6}$$

with s being the most distant source. In the case of just one lens and one source, Eq. 5 reduces to Eq. 1, with $\theta_2 = \beta$, $\theta_1 = \theta$, and $\eta_{ij} = 1$.

To reconstruct the unlensed sources in practice, the lens model should incorporate the deflections caused by s1 on s2. However, for the Cosmic Horseshoe system, the deflection contribution from s1 to the total angle experienced by s2 is expected to be negligible for three primary reasons: (i) the sources are at similar redshifts, so s1 will be an inefficient lens regardless of its mass (ii) The location of the radial image and its counter-image, imply that s1 is not closely aligned with s2 and that the images of s2 do not pass close to s1; and (iii) the low observed lensed surface brightness of s1, combined with its small size after reconstruction (see Section 4), suggests that it is a low-mass galaxy. Given these assumptions, in this work, we neglect the lensing effect of s1 on s2.

3.1.2 Lens modelling

For the lens modelling, we employ the open-source software `PyAutoLens` (Nightingale et al. 2018, 2021), which implements a Bayesian version (Suyu et al. 2006) of the semi-linear inversion (SLI)

method (Warren & Dye 2003). For a given set of non-linear parameters (describing the lens mass model and/or the source), the code linearly ray-traces image-pixels from the image plane back to the source plane, reconstructing the source emission using an adaptive mesh grid[8]

In PyAutoLens the *native* likelihood is that described in Suyu et al. (2006, eq. 19) and Dye et al. (2008, eq. 5). This approach incorporates the instrumental PSF blurring alongside regularisation terms for the pixelised source reconstruction, which helps to mitigate ill-posed solutions during the linear inversion. We refer the readers to these references for more details. [9].

To remove over- and under-magnified solutions (Maresca et al. 2021), we trace two image plane regions of s1 to the source plane, where they are expected to overlap. If these conjugate regions fail to overlap after delensing, we penalise the likelihood of this solution by a factor proportional to the distance between the two regions in the source plane. This approach effectively removes unphysical solutions resembling over/under-magnified versions of the data, without the need for parameter fine-tuning. Using regions, rather than pairs of conjugate points, also mitigates the risk of selecting incorrect pairs. This is particularly relevant in the case of the Horseshoe system, where the radial image is very faint and embedded within the lens light, making it challenging to reliably identify conjugate pairs between the radial image and its counter-image. Further details of this approach are provided in Appendix A.

In this work, we did not attempt to model the full Einstein ring of s2, as our primary focus is the inner mass distribution of the main deflector, which is better constrained by the lensed radial image of s1. However, the Einstein ring provides valuable information about the large-scale mass distribution, offering an additional constraint on the mass profile. Furthermore, the Einstein ring is sensitive to the mass of s1, as it reflects the total projected mass distribution along the LOS between the main-lens and s2.

To incorporate the lensing information from s2, we apply a Gaussian prior on the total mass enclosed within its Einstein radius, $M_{\rm Ein}$. This prior has a mean of $M_{\rm Ein} = 5.46 \times 10^{12} M_\odot$, corresponding to an effective Einstein radius of $R_{\rm Ein} = 5.08''$, and a standard deviation of $\sigma_{M_{\rm Ein}} = 0.27 \times 10^{12} M_\odot$. We derived $M_{\rm Ein}$ by modelling the lensed s2 source using an elliptical power-law (EPL) mass profile for the main deflector with PyAutoLens. The resulting values for $M_{\rm Ein}$ and $R_{\rm Ein}$ are consistent with those reported by Dye et al. (2008) and Schuldt et al. (2019). Fig. 3 shows the data pixels used to model the Cosmic Horseshoe, along with the highest-likelihood image model and the reconstructed source. Further details about the EPL mass model are provided in Appendix B.

It is important to highlight that the $M_{\rm Ein}$ prior accounts for the mass perturbation from s1. Specifically, $M_{\rm Ein}$ represents the total mass within a cylinder of radius $R_{\rm Ein}$, centred on the main-lens, along the LOS between the observer and s2. Nevertheless, as discussed in Section 4, excluding this prior from the likelihood does not impact the inferred BH mass.

3.2 Dynamical modelling

We described the dynamical state of the system using the Jeans formalism for a steady-state axisymmetric configuration (Binney & Tremaine 2008). In spherical coordinates (r, θ, ϕ), and assuming the velocity ellipsoid is aligned with the spherical coordinate system, the Jeans equations are (Cappellari 2020):

$$\frac{\partial(\nu\overline{v_r^2})}{\partial r} + \frac{2\nu\overline{v_r^2} - \nu\overline{v_\theta^2} - \nu\overline{v_\phi^2}}{r} = -\nu\frac{\partial\Phi}{\partial r},$$
$$\frac{\partial(\nu\overline{v_\theta^2})}{\partial\theta} + \frac{\nu\overline{v_\theta^2} - \nu\overline{v_\phi^2}}{\tan\theta} = -\nu\frac{\partial\Phi}{\partial\theta}, \tag{7}$$

where Φ is the total gravitational potential, (v_r, v_θ, v_ϕ) the velocities in spherical coordinates, and ν is the intrinsic luminosity density.

Defining the stellar anisotropy as

$$\beta_{\rm star} = 1 - \frac{\overline{v_\theta^2}}{\overline{v_r^2}} \equiv 1 - \frac{\sigma_\theta^2}{\sigma_r^2}, \tag{8}$$

and applying the boundary condition $\nu\overline{v_r^2} = 0$ as $r \to 0$, the Jeans equations 7 simplify to:

$$\nu\overline{v_\phi^2} = (1 - \beta_{\rm star})\left[\nu\overline{v_r^2} + \frac{\partial\left(\nu\overline{v_r^2}\right)}{\partial\theta}\tan\theta\right] + \nu\frac{\partial\Phi}{\partial\theta}\tan\theta, \tag{9}$$

and

$$\nu\overline{v_r^2} = \int_r^\infty \left(\frac{r'}{r}\right)^{2\beta_{\rm star}} \Psi(r', \theta')dr', \tag{10}$$

where

$$\theta' = \arcsin\left[\left(\frac{r'}{r}\right)^{\beta_{\rm star}-1}\sin\theta\right], \quad \text{and}$$
$$\Psi(r, \theta) = \nu\left(\frac{\partial\Phi}{\partial r} - \frac{\tan\theta}{dr}\frac{\partial\Phi}{\partial\theta}\right). \tag{11}$$

By solving Eqs. 9 and 10, the velocity moments can be integrated along the LOS (see, e.g., Section 3 on Cappellari 2020) to compute the observables, which can be compared to the galaxy's observed $v_{\rm rms}$ map.

For this purpose, we employed the Jeans Axisymmetric Multi-Gaussian Expansion (JAM) method (Cappellari 2008, 2020), as implemented in the Jampy software, to perform the stellar dynamical modelling. This approach assumes that the galaxy's mass distribution can be parameterised as a sum of concentric elliptical Gaussians (see Section 3.3).

For the dynamical modelling, we evaluated the goodness-of-fit using a χ^2-likelihood, which compares the observed data to the model predictions convolved with the MUSE PSF[10]. This ensures that the effects of seeing are properly accounted for in the analysis.

3.3 Multi-Gaussian Expansion (MGE)

The stellar light and mass component is modelled using the MGE method (Emsellem et al. 1994; Cappellari 2002), which parametrises

---

[8] This assumes that the lens light was subtracted already.

[9] Those interested in the technical aspects and implementation within PyAutoLens may also check the following notebooks: https://github.com/Jammy2211/autolens_workspace/tree/main/notebooks/advanced/log_likelihood_function

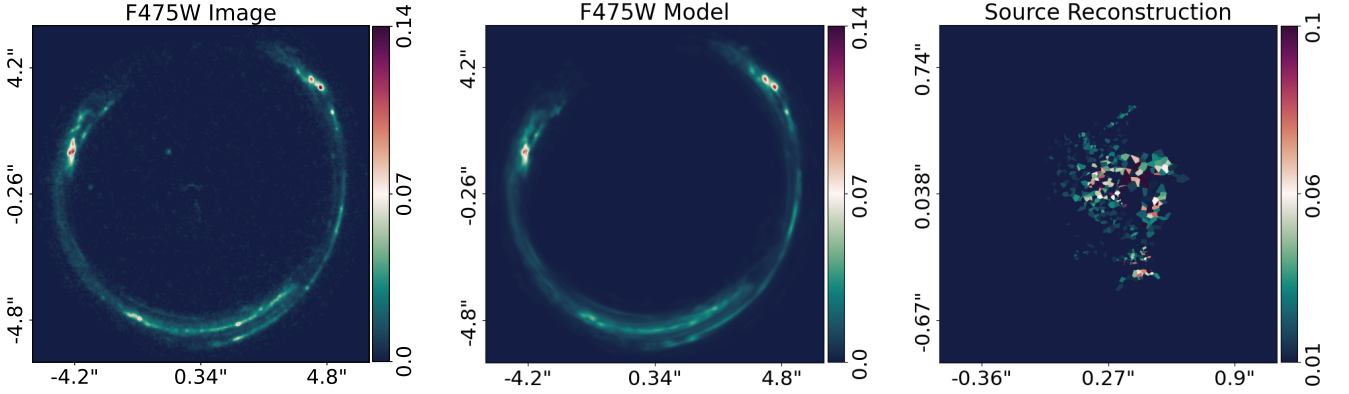[10] Here assumed to be the observed seeing.

Figure 3. Fits to the Cosmic Horseshoe in the F475W filter. From left to right, the panels display the lens-subtracted image, the highest-likelihood EPL model, and the reconstructed source s2 at $z = 2.381$. To enhance the efficiency of the lens modelling, we applied a mask around the lensed source and only modelled pixels within the masked region, as shown in the central panel. All images are in units of electrons per second. Additional details about the lens modelling are provided in Appendix B.

the stellar surface brightness as a sum of concentric elliptical Gaussians with the same orientation. If $I(x', y')$ represents the stellar surface brightness, its MGE parametrisation is given by

$$I(x', y') = \sum_{j=1}^{N} \frac{L_j}{2\pi\sigma_j^2 q_j'} \exp\left[-\frac{1}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{q_j'^2}\right)\right], \quad (12)$$

where N is the total number of Gaussians. The jth Gaussian component has a total luminosity L_j, an observed projected axial ratio $0 \leq q_j' \leq 1$, and a dispersion σ_j along the semi-major axis, aligned with x'-axis.

The three-dimensional luminosity density ν (and corresponding mass density) is obtained by deprojecting Eq. 12, assuming an inclination angle i, where $i = 90°$ corresponds to an edge-on orientation. The luminosity density ν is then converted into stellar mass density using a mass-to-light ratio, $\Upsilon_\star$.

Assuming an oblate axisymmetric model, the stellar mass density profile in cylindrical coordinates (R, ϕ, z) is given by (Cappellari 2002):

$$\rho(R, z) = \sum_{j=1}^{N} \frac{M_j}{(2\pi)^{3/2}\sigma_j^3 q_j} \exp\left[-\frac{1}{2\sigma_j^2}\left(R^2 + \frac{z^2}{q_j^2}\right)\right]. \quad (13)$$

Here, $M_j = \Upsilon_\star L_j$ represents the mass of the jth Gaussian component, with L_j luminosity and σ_j dispersion, as defined in Equation 12. The deprojected axial ratio, q_j, is given by

$$q_j^2 = \frac{q_j'^2 - \cos^2 i}{\sin^2 i}. \quad (14)$$

3.4 Mass profile

We described the mass profile of the main deflector using the MGE method, assuming a multicomponent mass model composed by (i) a stellar mass component; (ii) a DM halo component; and (iii) an additional central mass concentration representing a SMBH.

The stellar component was derived by deprojecting the observed surface brightness profile, as by Eq. 13. The DM halo follows a generalised Navarro-Frank-White (gNFW; Wyithe et al. 2001) profile,

$$\rho(r) = \rho_s \left(\frac{r}{r_s}\right)^{-\gamma_{\text{DM}}} \left(1 + \frac{r}{r_s}\right)^{\gamma_{\text{DM}}-3}, \quad (15)$$

where ρ_s is a characteristic density at the scale radius r_s, and γ_{DM} is the inner density slope that allows the profile to be cuspier ($\gamma_{\text{DM}} > 1$) or cored ($\gamma_{\text{DM}} = 0$). For $\gamma_{\text{DM}} = 1$, the profile reduces to the classical NFW (Navarro et al. 1997). To include this halo contribution in the dynamical model, we followed Cappellari et al. (2013) and also parametrised the DM component using the MGE method. Finally, the SMBH is modelled as an additional Gaussian with a small scale radius ($\sigma = 0.01''$).

The total mass distribution of the main-lens galaxy is then described by Eq. 13, with the summation over $N_{\text{star}} + N_{\text{DM}} + N_{\text{BH}}$ Gaussians, where N_j represents the number of Gaussians used to parameterise each respective mass component.

We also incorporate an external lensing shear to account for the perturbations of structures near the LOS, besides s1. The external shear field is parameterised by the two elliptical components ($\epsilon_1^{\text{sh}}, \epsilon_2^{\text{sh}}$). From these components, the shear magnitude γ^{sh} and shear angle ϕ^{sh}, measured counter-clockwise from north, are obtained as:

$$\gamma^{\text{sh}} = \sqrt{\epsilon_1^{\text{sh}2} + \epsilon_2^{\text{sh}2}}, \quad \tan\left(2\phi^{\text{sh}}\right) = \frac{\epsilon_2^{\text{sh}}}{\epsilon_1^{\text{sh}}}. \quad (16)$$

3.5 Joint modelling

Since lensing and dynamical data are independent, the joint likelihood is constructed as the product of their individual likelihood functions. In addition to these terms, the likelihood includes two additional terms: one enforcing consistency with the mass enclosed within the Einstein radius, as defined by the lensed image of s2 (M_{Ein}), and the penalty term defined in Sec. 3.1.2 that punishes solutions where s1 is significantly over- or under-magnified.

We break the modelling process into several stages, so as to make it more tractable

(i) Dynamical Model Fitting: We begin by fitting only the dynamical model, but with the addition of the prior term for M_{Ein} and the conjugate regions of s1. This provides an initial estimate of the mass parameters.

(ii) Source Grid Model of s1: Using the highest-likelihood dynamical model, we then sample the pixelised source grid of s1. A Voronoi grid is used to allocate more pixels to highly magnified regions of the source plane, along with a constant regularisation term. In `PyAutoLens`, this is represented by an `Overlay` mesh grid, with `VoronoiNN` pixelisation, and `ConstantSplit` regularisation.

(iii) Mass Parameter Resampling: After obtaining the source grid, we resample the mass parameters using the full likelihood, with s1 reconstructed based on the highest-likelihood model from the previous chain. This step improves the estimate of the mass parameters.

(iv) Source Plane Resampling: Using the highest-likelihood model, we resample the parameters of the source plane. The source is reconstructed in the source plane using a brightness-adaptive Voronoi mesh grid, with Natural Neighbour interpolation (Sibson 1981) and a regularisation scheme that adapts the degree of smoothing based on the source's surface brightness. This ensures that brighter regions are reconstructed at higher resolution, while regions with lower SNR are more regularised. In `PyAutoLens` notation, this corresponds to a `KMeans` mesh grid, with `VoronoiNN` pixelisation, and `AdaptiveBrightnessSplit` regularisation.

(v) Final Resampling of the Lens Macro Model: Finally, using the source grid that adapts to the source morphology, we sample the lens macro model once more, keeping the hyper-parameters of the source fixed at the highest-likelihood result.

Throughout all these stages, the same priors are applied across all chains for each given lens mass macro model. Further details are provided in Appendix C.

We sampled for the posterior distribution of the non-linear parameters using the nested sampler `dynesty` (Speagle 2020), a Python implementation of the nested sampling algorithm (Skilling 2006) designed for estimating Bayesian posteriors and evidences. The Bayesian evidence, $\mathcal{Z}$, which is the integral of the likelihood times the prior over the entire multidimensional parameter space, can be used to rank different models for the same dataset, providing an objective way to compare models (e.g., Liddle et al. 2006).

The Bayesian evidence naturally incorporates a penalty for increased model complexity – penalising models with additional free parameters, provides an objective basis for the principle of Occam's razor. Typically, when comparing models using the difference in their log-evidences, $\Delta \ln \mathcal{Z}$, a difference in $\Delta \ln \mathcal{Z} > 1$ is considered significant, > 2.5 strong, and > 5 is decisive, in favour of the model with the highest evidence.

4 RESULTS

To disentangle the baryonic matter from the DM contribution and to reveal the radial arc, a model for the lens light distribution is essential. We fitted the lens light in both HST/WFC3 filters, F814W and F475W, using the MGE method described in Section 3.3. During the fitting process, the radial and tangential lensed images were masked to minimise contamination from the source planes. The PSF effects were incorporated by modelling the `PSFr` model as a sum of circular Gaussians. The F814W lens light model was used to trace the stellar distribution, as it is a better tracer of the stellar mass budged and has less contamination from the radial arc. For the lens modelling, we utilised the F475W image after the lens light subtraction. The left panel of Fig. 3 shows the F475W lens-subtracted image, and Table 1 presents the MGE decomposition for the F814W band.

Hereafter, all parameter estimates are derived from the final sampling chain. Unless otherwise stated, values represent the median of each parameter's one-dimensional marginalised posterior probability

Table 1. MGE components of the HST/F814W image. The columns, from left to right, show the surface brightness of each Gaussian component, the MGE width, and the observed axis ratio. The MGE units were converted from counts to physical values following Trick et al. (2016), accounting for the redshift dimming effect and assuming a solar ST magnitude of $M_{\odot,\text{F814W}} = 5.35$ from Willmer (2018).

I [$L_\odot/\text{pc}^2$]	σ [arcsec]	q'
2416.68	0.0198	0.825
5883.51	0.1379	0.650
351.98	0.2026	1.000
1630.70	0.2161	0.650
770.51	0.3609	0.673
425.76	0.5039	1.000
171.75	1.1015	1.000
70.03	3.2367	1.000

distribution, with uncertainties corresponding to the 16$^{\text{th}}$ and 84$^{\text{th}}$ percentiles.

4.1 Fiducial model

The fiducial model (**M1**) is composed by a stellar mass component, an elliptical NFW[11] halo, and a central BH.

We assume that the stellar component follows the observed surface brightness distribution of the main-lens, scaled by a constant $\Upsilon_\star$. The DM component is assumed to be concentric with the baryonic matter and have the same alignment. Additionally, the scale radius r_s is fixed at 10 times the stellar effective radius, as seen in simulations (e.g., Kravtsov 2013) and used in other SGL studies (e.g., Sonnenfeld et al. 2015).

Beyond the mass parameters, the fiducial model includes four additional parameters: two elliptical components ($\epsilon_1^{\text{sh}}, \epsilon_2^{\text{sh}}$) that describe the external lensing shear; a constant stellar anisotropy parameter, β_{star}; and the inclination angle, i, of the lens relative to the LOS. Therefore, the fiducial model has eight parameters: $\left(i, \beta_{\text{star}}, \Upsilon_\star, \log_{10} \rho_s, q_{\text{DM}}, \log_{10} M_{\text{BH}}, \epsilon_1^{\text{sh}}, \epsilon_2^{\text{sh}}\right)$.

The median of the one-dimensional marginalised posterior of the fiducial model and their uncertainties are summarised in Table 2. Fig. 4 presents the two-dimensional posterior distributions for the BH mass and other parameters that exhibit significant degeneracies with it.

Our fiducial model favours a BH mass of $\log_{10}(M_{\text{BH}}/M_\odot) = 10.56^{+0.07}_{-0.08}$, with a notable degeneracy with the mass-to-light ratio and the DM characteristic density, as shown in Fig. 4. However, no significant degeneracy was observed between the BH mass and the anisotropy parameter. Similarly, the Einstein mass within the Cosmic Horseshoe, M_{Ein}, does not exhibit a strong degeneracy with the BH mass.
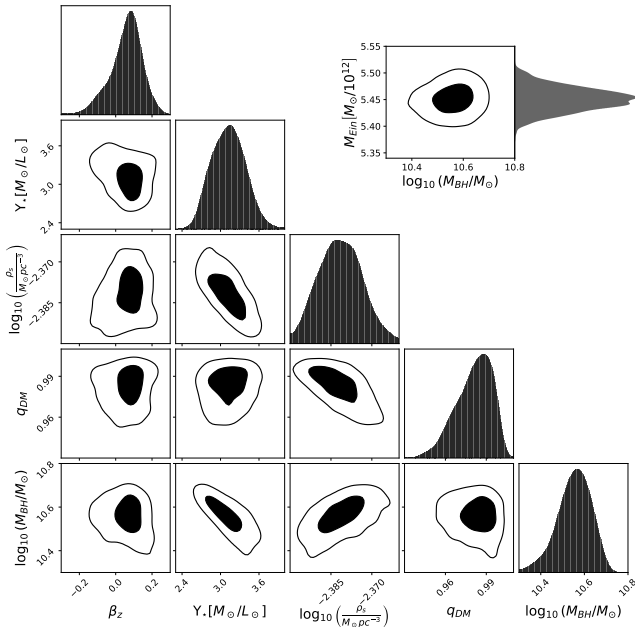
The median stellar mass-to-light ratio, $\Upsilon_\star = 3.13^{+0.25}_{-0.26} M_\odot/L_\odot$, is in agreement with values inferred from stellar population synthesis models of other ETGs (e.g., Conroy & van Dokkum 2012; Gu et al. 2022). This is a reasonable value for a massive elliptical galaxy: the expected value is $\sim 4 M_\odot/L_\odot$ for a simple stellar population with an age of 10Gyr and solar metallicity (Vazdekis et al. 2012; Ricciardelli et al. 2012; Verro et al. 2022).

We recovered a total projected mass within R_{Ein} of $M_{\text{Ein}} = 5.45^{+0.02}_{-0.03} \times 10^{12} M_\odot$, consistent with our prior and the EPL model,

---

[11] This is obtained by taking $\gamma_{\text{DM}} = 1$ in Eq. 15.

Table 2. Inferred median and 1σ credible intervals for the parameters of our fiducial model.

Parameter	Posterior (median with 1σ uncertainties)
$i\ [^\circ]$	65^{+15}_{-11}
$\beta_{\rm star}$	$0.07^{+0.06}_{-0.10}$
$\Upsilon_\star\ [M_\odot/L_\odot]$	$3.13^{+0.25}_{-0.26}$
$\log_{10}\left(\frac{\rho_s}{M_\odot {\rm pc}^{-3}}\right)$	$-2.38^{+0.01}_{-0.01}$
$q_{\rm DM}$	$0.98^{+0.01}_{-0.02}$
$\log_{10}(M_{\rm BH}/M_\odot)$	$10.56^{+0.07}_{-0.08}$
$\epsilon_1^{\rm sh}$	$0.01^{+0.01}_{-0.01}$
$\epsilon_2^{\rm sh}$	$0.05^{+0.01}_{-0.01}$
$M_{\rm Ein}\ [M_\odot/10^{12}]$	$5.45^{+0.02}_{-0.03}$

Figure 4. Two-dimensional posterior distributions for the parameters of the fiducial model. Only parameters that show a strong degeneracy with the BH mass are displayed. The inset plot on the top-right present the covariance between the BH mass and Einstein mass within the Cosmic Horseshoe ring. Contours are the 1 and 2σ credible intervals, respectively.

albeit slightly higher than the value reported by Schuldt et al. (2019). The DM fraction within R_e was found to be $f_{\rm DM}(\leq R_e) = 0.72^{+0.02}_{-0.02}$, which is also higher than Schuldt et al. (2019) and Spiniello et al. (2011), though it remains consistent with the 1σ range of the latter.

Fig. 5 shows the highest-likelihood lens model, alongside the normalised residuals, and the source plane reconstruction of source s1 at $z = 1.961$. Our model successfully reproduces the radial arc and its counter-image, while reconstructing the source's surface brightness, which exhibits an irregular morphology.

In Fig. 2, we present the results of the dynamical modelling. The top row displays the observed Voronoi-binned kinematic map, the median dynamical model, and the normalised residuals. Although the model exhibits a small squashing of the kinematic map along the semi-major axis, the radial profile (bottom panel) indicates a good fit to the data. Furthermore, the model was able to reproduce the

observed trend found in other BCGs (e.g., Smith et al. 2017b), where the kinematic profile flattens in the central regions and rises towards the edges.

4.2 Alternative mass models

To assess the robustness of our BH detection, we now investigate a range of alternative mass models. We first consider perturbations to our fiducial model and subsequently explore more flexible configurations. For each alternative model, we ran the full non-linear sampling again, following the steps outlined in section 3.5. Table 3 summarises the various mass models considered and their inferred BH masses, along with their corresponding Bayesian evidence relative to the fiducial model.

We present the medians and uncertainties of all posterior distributions in Appendix D.

4.2.1 Fiducial model perturbations

In this section, we investigate how perturbations around the fiducial model impact the outcomes. All $M_{\rm BH}$ from these model perturbations are summarised in Table 3.

4.2.1.1 gNFW profile: In our fiducial model **M1**, we considered an NFW halo for which the inner density slope is fixed. However, a steeper DM density profile, in the inner regions, could potentially compensate for the BH mass, leading to a more cuspy rather than cored DM distribution. To explore this possibility, we considered an alternative model (**M2**) in which the DM halo is parameterised by a gNFW profile, allowing the inner density slope, $\gamma_{\rm DM}$, to vary.

Fig. 6 shows the two-dimensional posterior distributions for the BH mass, the DM density slope, and the Einstein mass, with the median values of the fiducial model indicated by brown dashed lines. We can see that $\gamma_{\rm DM}$ exhibits only a modest degeneracy with the BH mass and a marginal degeneracy with the Einstein mass. Furthermore, the DM inner density slope was found to be $\gamma_{\rm DM} = 1.06^{+0.05}_{-0.07}$, in agreement with an NFW profile. The BH mass is consistent with the fiducial model within 2σ.

4.2.1.2 Variable anisotropy profile: Next, we explored the possibility of a variable anisotropy profile. As shown by Thomas et al. (2014), massive ETGs hosting SMBHs often exhibit a radial variation in the anisotropy parameter, $\beta_{\rm star}$. In particular, for core galaxies, stellar motions within the core radius tend to be dominated by tangential orbits ($\beta_{\rm star} < 0$), while outside the core, the orbits become more radially dominated ($\beta_{\rm star} > 0$).

To assess the impact of stellar anisotropy on the BH mass estimate, we introduced a stellar anisotropy profile, $\beta_{\rm star}(r)$, in model **M3**. This profile is constructed using the luminous MGE components (Cappellari 2008) to define regions with distinct anisotropy parameters. We used the Gaussian width as a proxy for the radius of influence of the parameter. Specifically:

- Components with $\sigma \leq 0.1''$ are assigned an anisotropy $\beta_{\rm star}^0$,
- Components with $0.1 < \sigma \leq 1.0''$ are assigned $\beta_{\rm star}^1$, and
- Components with $\sigma \geq 1.0''$ are assigned $\beta_{\rm star}^2$.

Furthermore, we allowed each $\beta_{\rm star}$ to be independent of each other. The rest of the mass model in **M3** remains identical to that of the fiducial model.

In Fig. 7, we present the stellar anisotropy profile for model **M3**, calculated following Cappellari (2008). Unlike other BCGs hosting
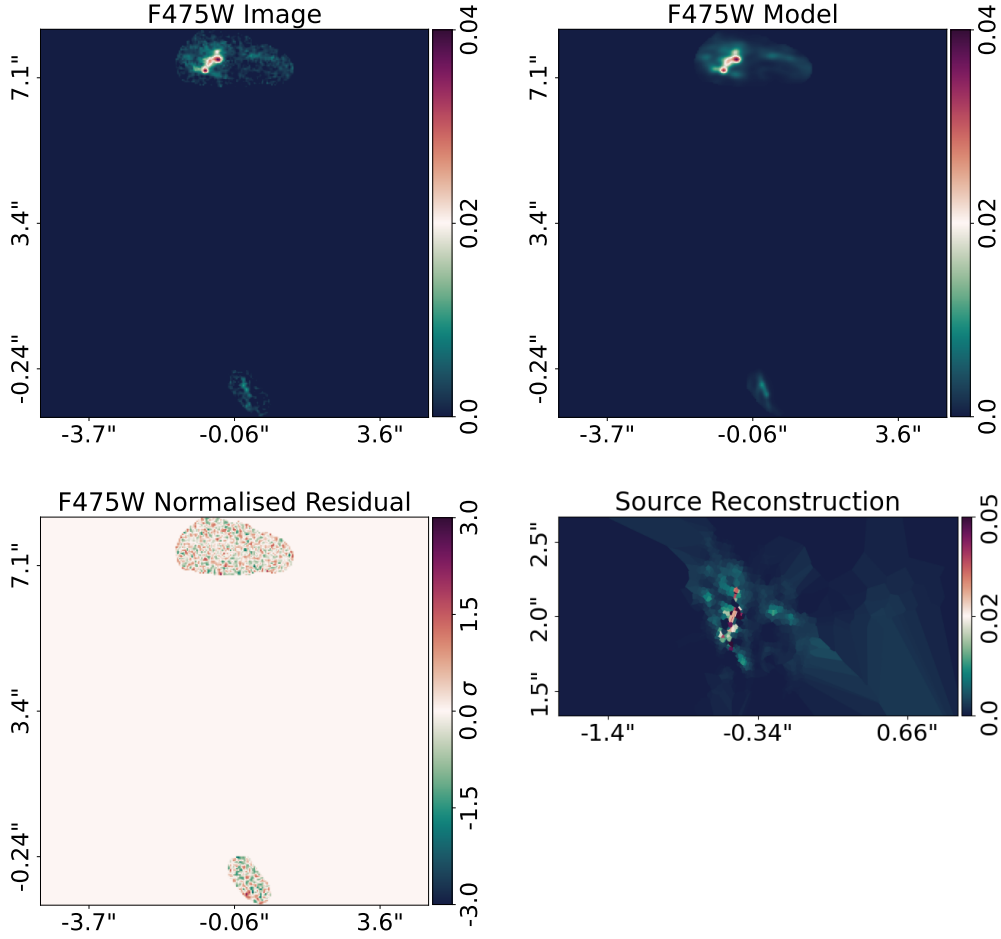
Figure 5. Highest-likelihood lens model under the fiducial configuration. Panels, from top left to bottom right, show the observed image, the lensed source model, the normalised residuals, and the source reconstruction. All images are in units of electrons per second. **There is also an interpolated version of the source, if preferred.**

central SMBHs, the anisotropy profile of our system near the galaxy centre is qualitatively distinct, showing no dominance of tangential orbits. (e.g., Thomas et al. 2014; Mehrgan et al. 2019). It is worth noting, however, that we cannot resolve the galaxy core (if present), where tangential orbits are typically expected. Moreover, the uncertainties near the centre suggest that $\beta_{\rm star}$ could also assume negative values in this region. Despite that, the radial variation observed in our profile is broadly consistent with findings by Gerhard et al. (2001) in elliptical galaxies and with simulations of massive ETGs (Wu et al. 2014). The BH mass derived from this model is slightly lower than the value obtained from the fiducial model but remains within the 1σ confidence level.

4.2.1.3 Varying mass-to-light ratio: Another way to account for the BH mass is by increasing the stellar mass at the galaxy's centre, representing an excess mass linked to a gradient in the stellar mass-to-light ratio (e.g., Smith et al. 2017a). To test this hypothesis, in model **M4**, we allowed the mass-to-light ratio to vary across different sets of Gaussian components. As in model **M3**, we defined the mass-to-light ratio for each luminous Gaussian component based on the Gaussian width, applying the same width constraints. This setup resulted in three distinct mass-to-light ratios ($\Upsilon_\star^0, \Upsilon_\star^1, \Upsilon_\star^2$), which were

constrained during the non-linear sampling to follow a decreasing gradient with radius.

In Fig. 8, we present the two-dimensional posterior distributions for the BH mass and the three mass-to-light ratios in model **M4**. The top-right inset plot compares the mass-to-light ratio profiles of **M1** and **M4**. The gradient profile of model **M4** is shown in black, while the constant profile of the fiducial model is shown in brown. Although model **M4** suggests a radial gradient, the large uncertainties make it consistent with the constant value recovered by the fiducial model. By this figure, it is clear the strong degeneracy between these parameters. Despite that, the recovered BH mass is consistent with the fiducial model within 1σ.

4.2.1.4 DM scale radius: In the fiducial model, we assume a DM scale radius fixed at ten times the galaxy's effective radius, and although motivated by simulations, such rigid constraints could bias the results. To evaluate the impact of this assumption, in model **M5** we allowed the scale radius to vary. This model recovered a scale radius of $r_s = 19.16^{+2.30}_{-0.97}$, interestingly consistent with the fiducial model's assumption. Moreover, the BH mass remains consistent with the fiducial value within 1σ.

Table 3. Joint fiducial and alternative mass models. The columns, from left to right, are: model name identification, mass model configuration, median of the BH mass posterior distribution, and the difference in the natural logarithm of the Bayesian evidences relative to the fiducial model (**M1**). Superscripts on the left side means the number of free components, i.e., $^3\beta_{\rm star}$ implies three free anisotropy parameters.

Model ID	Mass model	$\log_{10}(M_{\rm BH}/M_\odot)$	$\Delta \ln \mathcal{Z}$
M1	$\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$ + BH	$10.56^{+0.07}_{-0.08}$	0.00
M2	$\Upsilon_\star$ + ell. gNFW ($r_s = 10R_e$) + $\beta_{\rm star}$ + BH	$10.57^{+0.07}_{-0.09}$	-0.53
M3	$\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $^3\beta_{\rm star}$ + BH	$10.45^{+0.11}_{-0.14}$	-3.48
M4	$^3\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$ + BH	$10.53^{+0.10}_{-0.11}$	2.01
M5	$\Upsilon_\star$ + ell. NFW + $\beta_{\rm star}$ + BH	$10.56^{+0.08}_{-0.08}$	-0.38
M6	$\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$ + BH w/ cyl. velocity ellipsoids	$10.55^{+0.08}_{-0.09}$	-0.36
M7	$\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$ + BH w/ Delaunay pixelisation	$10.55^{+0.08}_{-0.08}$	-2.27
M8	$\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$ + BH w/o Horseshoe	$10.51^{+0.07}_{-0.09}$	0.38
M9	$^3\Upsilon_\star$ + ell. gNFW ($r_s = 10R_e$) + $^3\beta_{\rm star}$ + BH	$10.50^{+0.10}_{-0.32}$	8.29
M10	Gaussian $\Upsilon_\star$ + sph. NFW ($r_s = 10R_e$) + $^8\beta_{\rm star}$ + BH	$10.55^{+0.10}_{-0.07}$	1.17
M11	$^3\Upsilon_\star$ + ell. gNFW w/ main $c(M, z)$ + $^3\beta_{\rm star}$ + BH	$10.15^{+0.17}_{-0.30}$	2.92
M12	$^3\Upsilon_\star$ + ell. gNFW w/ 1σ bellow $c(M, z)$ + $^3\beta_{\rm star}$ + BH	$10.33^{+0.07}_{-0.13}$	10.73
M13	$^3\Upsilon_\star$ + ell. gNFW w/ 1σ above $c(M, z)$ + $^3\beta_{\rm star}$ + BH	$10.59^{+0.04}_{-0.10}$	11.48
M14	$\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$	-	16.31
M15	$^3\Upsilon_\star$ + ell. NFW ($r_s = 10R_e$) + $\beta_{\rm star}$	-	11.17

4.2.1.5 Modelling choices: There are three other sources of systematics linked to the fiducial model. First, the alignment of the velocity ellipsoid, which is assumed to align with the spherical coordinate system. Second, the source plane pixelisation grid, which is based on a Voronoi tessellation. And third, the use of the prior on the total projected mass within the Horseshoe.

In model **M6**, we test the effect of a cylindrical alignment (Cappellari 2008) for the velocity ellipsoid, while in model **M7**, we change the source plane pixelisation from a Voronoi to a Delaunay tessellation. In both cases, we recover BH masses that agree with the fiducial model within the 1σ level.

Finally, we assess the impact of the prior on the total projected mass within the Einstein radius of s2 by removing this constraint in model **M8**. This model infers a BH mass of $\log_{10}(M_{\rm BH}/M_\odot) = 10.51^{+0.07}_{-0.09}$, showing that the prior has a minimum impact on the determination of the BH mass.

4.2.2 More flexible mass models

We now explore more flexible mass models to assess whether increased freedom in the mass distribution can account for the high BH mass inferred in our fiducial model.

In model **M9** we assumed a gradient mass-to-light ratio, defined the same way as in model **M4**. Additionally, we considered an anisotropy profile with three independents anisotropy parameters, as in model **M3**. We also considered an gNFW profile for the halo mass, and as before we kept the scale radius fixed at ten times R_e. For model **M9**, we found $\log_{10}(M_{\rm BH}/M_\odot) = 10.50^{+0.10}_{-0.32}$. For this model, we also recovered a DM inner slope of $\gamma_{\rm DM} = 1.08^{+0.06}_{-0.07}$, which still consistent with a NFW profile.

In model **M10**, we introduced a more flexible mass-to-light ratio

profile by parametrising it as a Gaussian-modulated function:

$$\Upsilon_\star^j = \Upsilon_0 \left[\upsilon_0 + (1 - \upsilon_0)e^{-0.5(\sigma_j \, \delta)^2} \right], \qquad (17)$$

where Υ_0 is the central stellar mass-to-light ratio, δ is a gradient parameter describing the profile's smoothness, υ_0 is the ratio between the central and outermost values, and σ_j represents the dispersion of the $j^{\rm th}$ MGE component. This approach enables each luminous Gaussian to have its unique $\Upsilon_\star^j$, while maintaining a small number of free parameters ($\Upsilon_0, \upsilon_0, \delta$) and ensuring a naturally decreasing profile.

Further, model **M10** incorporates additional freedom by assigning each luminous Gaussian its own anisotropy parameter. The DM halo is modelled as a spherical NFW profile with the scale radius fixed to the fiducial model. With this configuration, we recover a BH mass of $\log_{10}(M_{\rm BH}/M_\odot) = 10.55^{+0.10}_{-0.07}$.

The final set of models explores the impact of adopting the mass-concentration relation from Ludlow et al. (2016) to define the DM scale radius of a gNFW profile. In these models, the DM characteristic density is parameterised by the mass at 200 times the critical density of the Universe, $M_{200}^{\rm DM}$, which is treated as a free parameter. Using $M_{200}^{\rm DM}$, the DM scale radius is derived based on the main relation from Ludlow et al. (2016), as well as the 1σ scatter above and below it. The anisotropy profile is set as in model **M3**, while the mass-to-light ratio is parameterised as in model **M4**. For these models, we found that:

- Model **M11** assuming the main mass-concentration relation, yields a BH mass of $\log_{10}(M_{\rm BH}/M_\odot) = 10.15^{+0.17}_{-0.30}$.
- Model **M12** applies the 1σ below the mean relation, resulting in $\log_{10}(M_{\rm BH}/M_\odot) = 10.33^{+0.07}_{-0.13}$.
- Model **M13** adopts the 1σ above the mean relation, recovering $\log_{10}(M_{\rm BH}/M_\odot) = 10.59^{+0.04}_{-0.10}$.
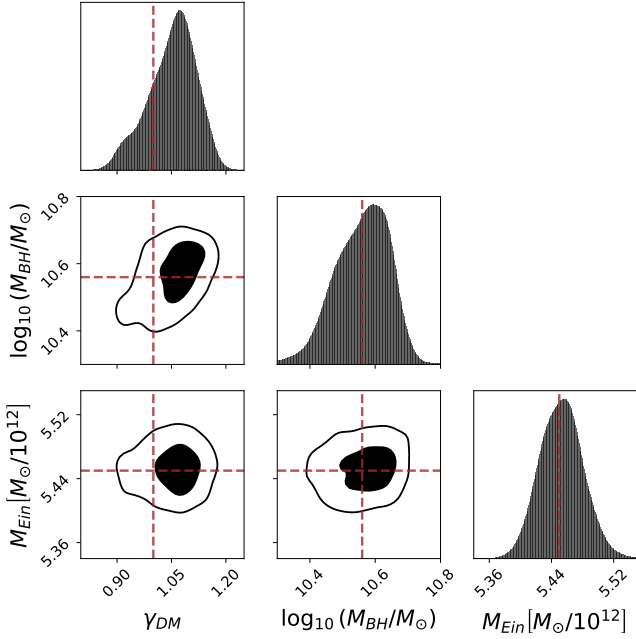
Figure 6. Two-dimensional posterior distributions for the parameters of model **M2**, which differs from the fiducial model by assuming a gNFW halo. Only the inner DM density slope, the BH mass, and the Einstein mass are displayed. The brown dashed lines shows the posterior median of the fiducial model for comparison. Contours are the 1 and 2σ credible intervals, respectively.
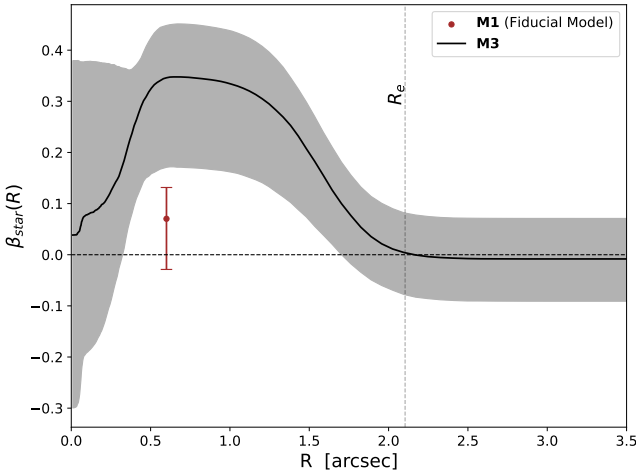

Figure 7. M3 anisotropy profile. The solid line shows the orbital anisotropy profile of the stars, and the gray band represents the 1σ credible band. The brown point is the constant anisotropy inferred by the fiducial model **M1**, with its associated error bar. The horizontal dashed line corresponds to the isotropic case.

These results, once more, highlight the robustness of the fiducial model.

4.3 Is a SMBH necessary?

So far, we only fit models with a presence of a SMBH, and despite the many variations of the mass profile, our results are fairly consistent between each other. But one could ask if it is necessary a BH to fit the
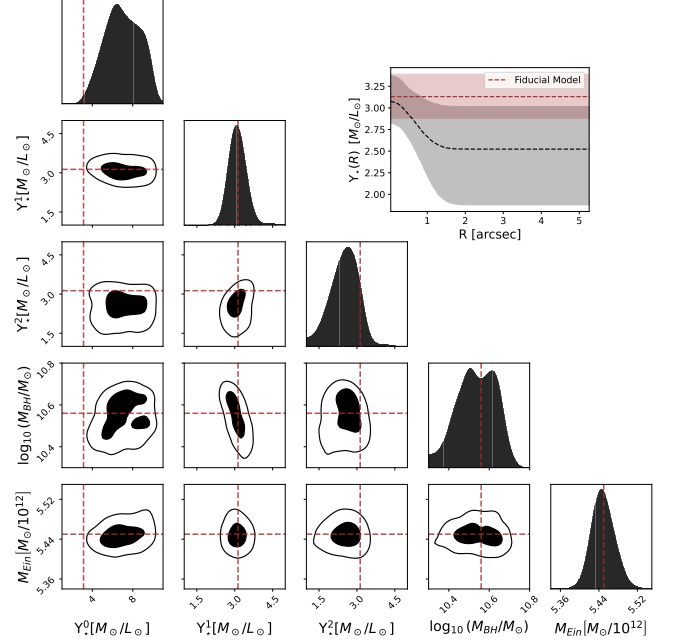

Figure 8. Two-dimensional posterior distributions for the parameters of model **M4**, which differs from the fiducial model by a gradient mass-to-light ratio. Only the three mass-to-light ratios, the BH mass, and the Einstein mass are displayed. The brown dashed lines shows the posterior median of the fiducial model for comparison. Contours are the 1 and 2σ credible intervals, respectively. The top-right inset shows the projected radial stellar mass-to-light profiles for the fiducial model (brown) and model **M4** in black. The shaded regions are the 1σ level.

data. To answer this question, we fitted model **M14** using the same fiducial mass model, but without including the BH component.

Fig. 9 shows the highest-likelihood lens model in the upper panels, and the dynamical model in the bottom panels for model **M14**. Qualitatively, the fiducial lens model and the **M14** lens model are the same. Both are able to reproduce the observed data with similar residuals, and to reconstruct the source with similar morphologies. On the other hand, the dynamical model fit the data poorly, especially at the central regions, where the BH is expected to be more relevant.

We can also use the Bayesian evidence of the models to quantitatively assess the need of a SMBH. The last column in Table 3 summarises the Bayesian evidence, $\ln \mathcal{Z}$, for the models considered in this work. Comparing the fiducial model **M1** and model **M14**, which excludes the SMBH, we find a difference in the Bayesian evidence of $\Delta \ln \mathcal{Z} = 16.31$. This corresponds to a statistical preference exceeding 5σ in favour of the SMBH[12].

We also attempt to fit model **M4** without the inclusion of an SMBH (**M15**) to evaluate whether the central mass-to-light component could replicate the BH's contribution. In this configuration, the central Gaussian component is assigned its own $\Upsilon_\star$, potentially steepening the stellar mass density profile near the galaxy's centre and compensating for the absence of the SMBH. However, as with model **M14**, this approach failed to reproduce the observed kinematical data accurately. A comparison of the Bayesian evidence for this model against the fiducial model yields $\Delta \ln \mathcal{Z} = 11.17$, strongly favouring the fiducial model. This difference corresponds to a 5σ detection of the SMBH.

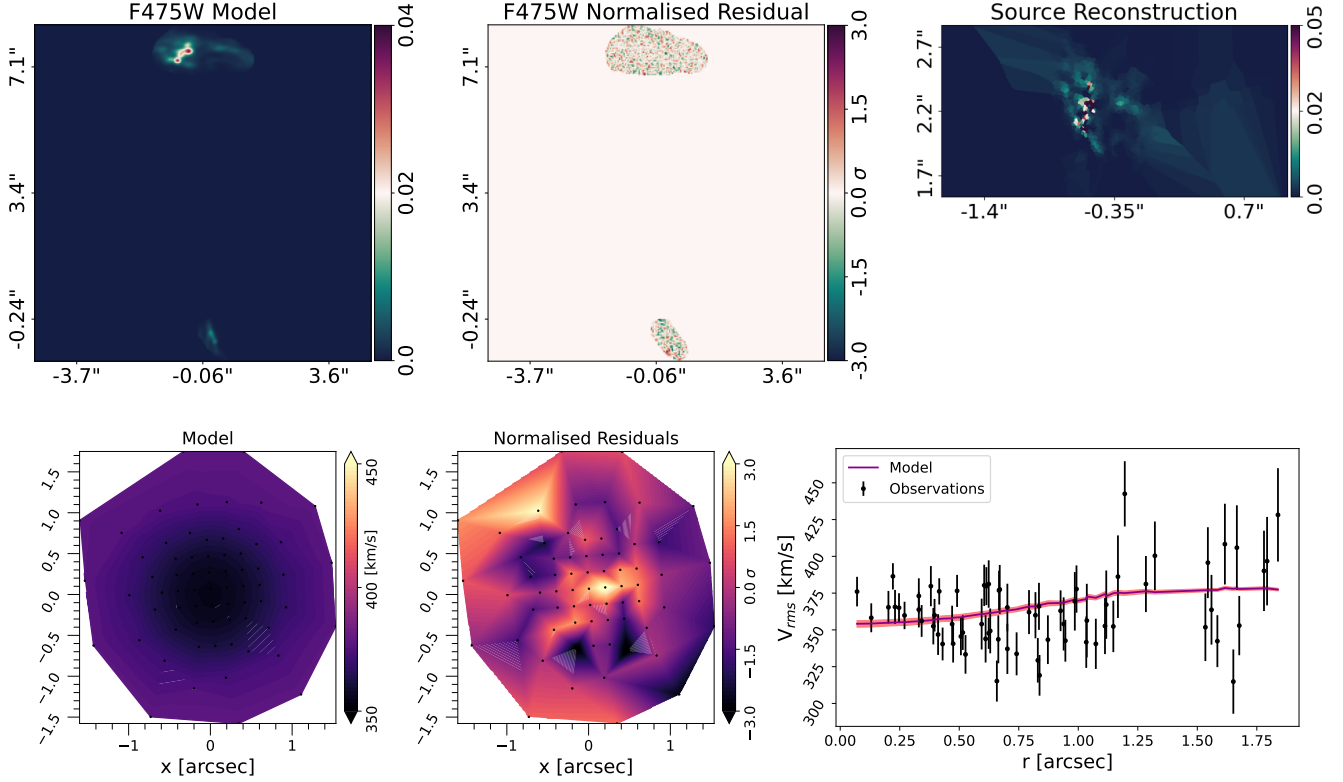[12] This assumes equal prior model probabilities.

Figure 9. Model results of mass configuration **M14**, which differs from the fiducial model by not including a SMBH. The upper panels display, left to right, the highest-likelihood lens model, the normalised residuals, and the source reconstruction. The bottom panels are, left to right, the median dynamical model, the normalised residuals, and the radial kinematic profile with the median model. All top images are in units of electrons per second.

4.4 The SMBH mass

In Table 3, we present the BH masses and Bayesian evidences for all the mass models in this work. The highest Bayesian evidence is associated with model **M3**, which incorporates a variable anisotropy profile. A comparison between the fiducial model **M1** and model **M3** yields $\Delta \ln \mathcal{Z} = -3.48$, corresponding to a 0.9σ preference to **M3**. This evidence difference is not decisive, and we do not have sufficient prior knowledge to say if one model should be astrophyscially preferred over the other. Additionally, measuring the anisotropy profile is notoriously challenging due to its degeneracies with other parameters and sensitivity to data quality, which is why we adopted the fiducial model for its simplicity and robustness.

Given this result, we adopt the BH mass inferred by the fiducial model as our final value, and we take the scatter between the alternative mass as an estimate of the systematic uncertainty. Using the standard deviation across all BH mass measurements, our final inference is $\log_{10}(M_{\mathrm{BH}}/M_\odot) = 10.56^{+0.07}_{-0.08} \pm (0.12)^{\mathrm{sys}}$ at 1σ level, confirming the detection of an UMBH in the Cosmic Horseshoe main-lens galaxy.

4.5 The role of the radial image

As we saw when comparing models **M1** and **M14**, the absence of the SMBH has a relatively modest effect on the lens model, but significantly impacts the fit to the kinematical data. This naturally raises the question of the role that lensing information plays in determining the SMBH mass in this case.

To explore this, we performed dynamical-only modelling for all

Table 4. Dynamical models only BH results. Values are the median and 1σ uncertainties. Models **M7** and **M8** are not applicable.

Model ID	$\log_{10}(M_{\mathrm{BH}}/M_\odot)$	Model ID	$\log_{10}(M_{\mathrm{BH}}/M_\odot)$
M1	$10.72^{+0.10}_{-0.13}$	**M9**	$10.19^{+0.50}_{-1.69}$
M2	$10.70^{+0.13}_{-0.17}$	**M10**	$9.78^{+0.65}_{-1.16}$
M3	$10.41^{+0.16}_{-0.62}$	**M11**	$10.30^{+0.35}_{-0.99}$
M4	$10.69^{+0.10}_{-0.15}$	**M12**	$10.59^{+0.19}_{-0.69}$
M5	$10.77^{+0.08}_{-0.11}$	**M13**	$10.35^{+0.25}_{-1.30}$
M6	$10.79^{+0.09}_{-0.12}$		

the mass models listed in Table 3, and we show the resulting BH mass measurements in Table 4.

Comparing the jointly results with those obtained through dynamical-only modelling, we find that the latter generally yields more massive BH estimates. Additionally, the error bars for the dynamical-only models are larger, which is expected due to the smaller number of data points in the kinematic map. Using the same criteria as before to determine the final BH mass, the dynamical-only BH mass is $\log_{10}(M_{\mathrm{BH}}/M_\odot) = 10.72^{+0.10}_{-0.13} \pm (0.30)^{\mathrm{sys}}$ at 1σ level.

These findings highlight the important role of lensing information in constraining the BH mass, particularly by limiting how massive the BH can be. This is especially significant in the context of direct SMBH mass determinations in intermediate and high-redshift systems, where IFU data often suffers from suboptimal spatial resolution and SNR. When the radial image is well-resolved and has
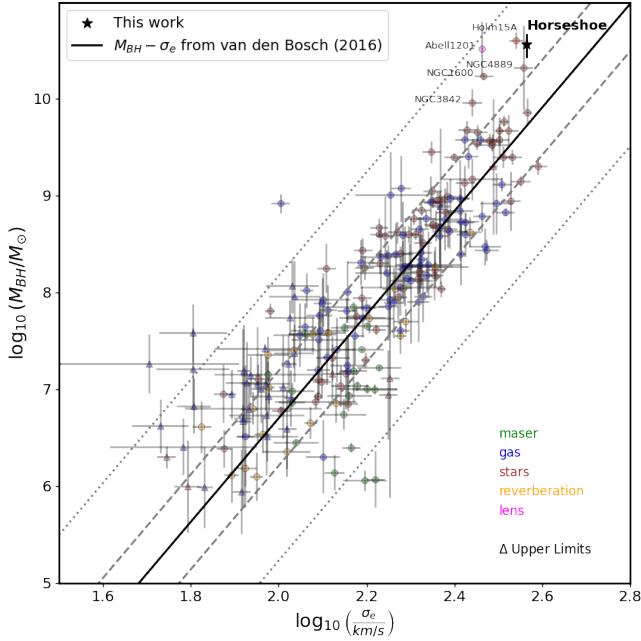
Figure 10. Relationship between the SMBH mass and the host effective velocity dispersion. The black solid line represents the relation from [van den Bosch (2016)](#), with dashed and dotted lines showing the 1σ and 3σ scatter, respectively. The UMBH at the centre of the Cosmic Horseshoe's main lens is marked by a star, with a measured mass of $\log_{10}(M_{BH}/M_\odot) = 10.56^{+0.07}_{-0.08} \pm (0.12)^{\rm sys}$. Other UMBHs that deviate significantly from the $M_{\rm BH} - \sigma_e$ relation are also shown: NGC 4889 and NGC 3842 ([McConnell et al. 2011](#)), NGC 1601 ([Thomas et al. 2016](#)), Holm 15A ([Mehrgan et al. 2019](#)), and Abell 1201 ([Nightingale et al. 2023](#)). These systems are typically BGCs, and except Abell 1201 at $z = 0.169$, they are all nearby systems. The Cosmic Horseshoe, at $z_l = 0.44$, represents one of the most massive SMBHs measured and is an $\sim 1.5\sigma$ outlier from the main $M_{\rm BH} - \sigma_e$ relation.

sufficient SNR, the lensing effect is sufficient to effectively constrain the BH mass, as demonstrated by [Nightingale et al. (2023)](#). On the other hand, when image quality is less favorable — reflecting the challenges of observing more distant systems — the integration of dynamical and lensing data becomes essential for reliable SMBH mass measurements.

5 DISCUSSION

5.1 SMBHs and the $M_{\rm BH} - \sigma_e$ relation

In Fig. [10](#), we put the SMBH of the main deflector in the Cosmic Horseshoe lens system in the context of the $M_{\rm BH} - \sigma_e$ relation from [van den Bosch (2016)](#). The SMBH reported here is among the most massive BH ever detected, but so is the galaxy that hosts it: the measured effective velocity dispersion of $\sigma_e = 366 \pm 6 \, \rm km \, s^{-1}$. Other UMBHs, with similar σ_e, are Holm 15A ([Mehrgan et al. 2019](#)), and NGC4889 ([McConnell et al. 2011](#)), both are BGCs and nearby galaxies. The lens of the Horseshoe is unique in that is at $z = 0.44$ and that has no comparably massive companion galaxies — it is likely a fossil group ([Ponman et al. 1994](#)).

Considering the $M_{\rm BH} - \sigma_e$ relationship from [van den Bosch (2016)](#), the SMBH we measured is an $\sim 1.5\sigma$ outlier, appearing overly massive for the host galaxy's effective velocity dispersion. In fact, the very high-mass end of the $M_{\rm BH} - \sigma_e$, predominantly populated by BCGs, shows this distinct trend, with SMBH masses systematically

exceeding the mean relation ([Bogdán et al. 2018](#)). This deviation at the massive end likely reflects distinct evolutionary pathways during the formation and assembly of these galaxies.

One possible scenario involves binary SMBH scouring, a process that can occur during the merger of massive galaxies and is more likely in the central galaxy of a group or cluster. In this process, the binary SMBHs dynamically expel stars from the central regions of the merged galaxy, effectively reducing the stellar velocity dispersion while leaving the SMBH mass largely unchanged (e.g., [Thomas et al. 2014](#), [2016](#); [Dullo 2019](#)). Another possible scenario involves AGN feedback processes, where powerful outflows and jets may quench star formation and alter the galaxy's central structure, decoupling the growth of the SMBH from the host galaxy's stellar kinematics. Strong AGN feedback can also transfer energy to the DM and stellar components, modifying the central surface brightness profile and mimicking the presence of a core (see discussion in [Mehrgan et al. 2019](#)). A third scenario posits that such UMBH could be remnants of extremely luminous quasars, which experienced rapid SMBH accretion episodes in the early Universe ([McConnell et al. 2011](#); [Wu et al. 2015](#)).

These distinct mechanisms highlight the complexity of galaxy and SMBH co-evolution, particularly for the most massive galaxies, and underscore the need for tailored models (and further observations) to explain the scatter in the $M_{\rm BH} - \sigma_e$ relation at its upper end.

5.1.1 Other astrophysical implications

Beyond the determination of the BH mass, we can infer other physical properties of the main deflector.

As discussed, while model **M4** suggests a gradient in the mass-to-light ratio, the constant value inferred by the fiducial model remains consistent with it within the uncertainties. The fiducial model predicts a projected stellar mass fraction within the Einstein radius of $f_\star (\leq R_{\rm Ein}) = 0.13^{+0.01}_{-0.01}$, which agrees with the value reported by [Spiniello et al. (2011)](#), supporting a Salpeter initial mass function. Similarly, model **M4** gives $f_\star (\leq R_{\rm Ein}) = 0.11^{+0.02}_{-0.02}$, also consistent with the previous findings. Even under the more flexible assumptions of model **M10**, where the mass-to-light ratio is modulated by a Gaussian function, the projected stellar mass fraction remains in agreement with the fiducial result, at $f_\star (\leq R_{\rm Ein}) = 0.10^{+0.01}_{-0.01}$.

The inner DM density slope is another noteworthy quantity, as it provides critical insights into the interaction between baryons and DM (e.g., [Gnedin et al. 2004](#); [Petit et al. 2023](#)). Early N-body DM-only simulations suggested that haloes are well described by the NFW profile ([Navarro et al. 1997](#)). However, the inclusion of baryonic components, especially feedback processes (e.g., [Di Cintio et al. 2014](#); [Jackson et al. 2024](#)), has been shown to alter the DM distribution within galaxies. These modifications may be linked to longstanding issues such as the "cusp-core" problem (see [Del Popolo & Le Delliou 2022](#), for a review).

In our analysis, while the fiducial model assumes an NFW halo, we introduced more flexibility in the inner DM density slope through models **M2** and **M9**, both of which assume a gNFW halo. For model **M2**, we obtained an inner DM slope of $\gamma_{\rm DM} = 1.06^{+0.05}_{-0.07}$, and for model **M9**, $\gamma_{\rm DM} = 1.08^{+0.06}_{-0.07}$. Both results are consistent with an NFW-like halo.

Fig. [11](#) compares the surface mass density profiles along the semi-major axis for these three models. All models exhibit strong agreement, particularly in the inner regions. The most notable deviation occurs in the outermost region of the stellar density profile for model **M9**, but this remains within the 1σ uncertainty. The larger uncer-
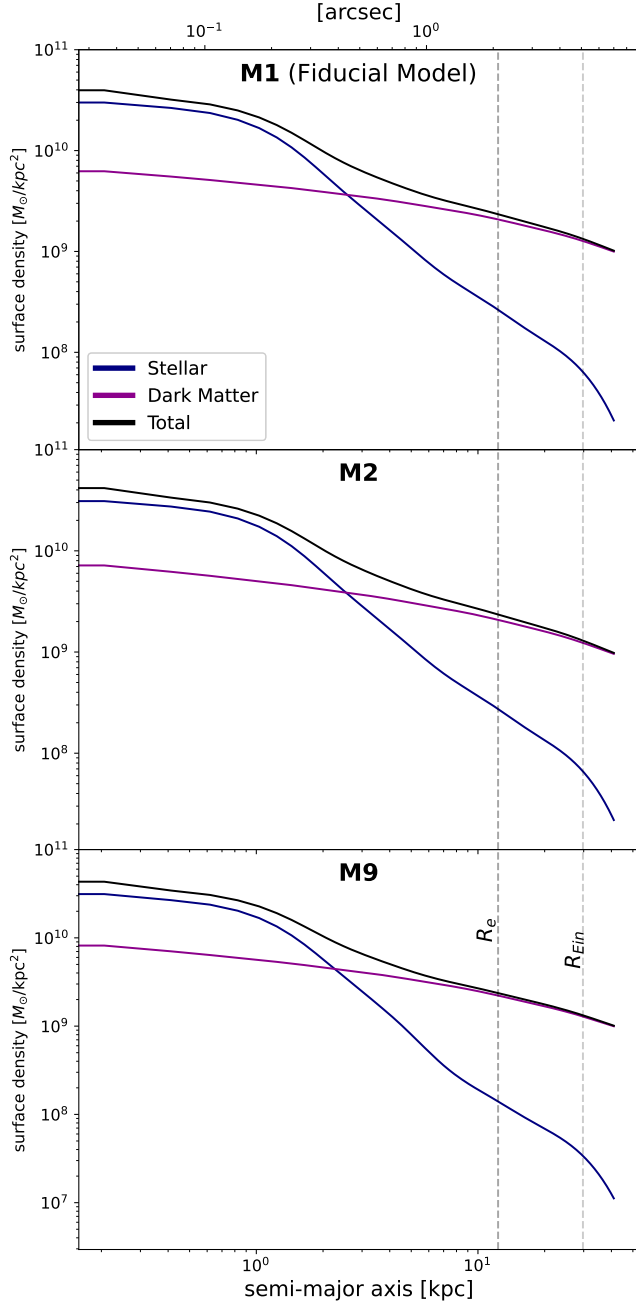
Figure 11. Surface mass density profile for three models — fiducial (upper) and other two that uses an gNFW profile: **M2** (middle) and **M9** (lower). Blue lines represent the stellar component, purple lines represents the DM component, and in black we show the total (DM+stellar+BH) surface density. The shaded regions indicate the 1σ credible intervals for each component. The horizontal dashed lines represents the effective radius and the Einstein radius.

tanties in model **M9** can be explained by the greater freedom in its mass profile, which allows for simultaneous variations in the stellar mass-to-light ratio, stellar anisotropy, and DM inner slope. Additionally, the galaxy is already DM dominated before reaching the effective radius

We also employed a gNFW halo in models **M11**-**M13**, but incorporating a mass-concentration relation to determine the scale radius. The recovered inner DM slopes for these models were

$\gamma_{DM} = 1.29^{+0.05}_{-0.04}$, $\gamma_{DM} = 1.34^{+0.07}_{-0.07}$, and $\gamma_{DM} = 1.12^{+0.03}_{-0.04}$ respectively. These slopes are steeper than an NFW-like profile. However, it is important to note that these steeper slopes do not result in better fits or higher Bayesian evidences. On the contrary, models **M11**-**M13** exhibit lower evidences compared to the other gNFW models that recovered an NFW-like slope. Additionally, when comparing these models with the fiducial model, the fiducial model remains slightly preferred.

This topic was also investigated by Schuldt et al. (2019) in their analysis of the Cosmic Horseshoe. They reached a similar conclusion, finding that allowing more flexibility in the inner density slope of the DM halo did not lead to a significant improvement in the fit. Their results, like ours, suggest that the halo is either NFW-like or deviates only slightly from this profile, indicating that baryonic processes either do not significantly alter the inner DM halo structure in this galaxy, or cancel each other out.

6 CONCLUSIONS

In this work, we investigated the inner structure of the unique Cosmic Horseshoe lensing galaxy by applying a self-consistent model to both the radial arc and the stellar dynamics of the main deflector, using the mass of the main ring as a prior. Our fiducial model constrained the mass of the UMBH at the centre of the main-lens to be $\log_{10}(M_{BH}/M_\odot) = 10.56^{+0.07}_{-0.08} \pm (0.12)^{\text{sys}}$. We rigorously tested a variety of systematics, including uncertainties in the mass profile and modeling choices, but all models consistently converged to the fiducial value. A Bayesian model comparison revealed a 5σ detection of the UMBH relative to a model without a BH contribution, reinforcing our results.

This mass places the Cosmic Horseshoe $\sim1.5\sigma$ above the $M_{BH} - \sigma_e$ relation (Fig. 10), may suggesting a unique evolutionary history for the Cosmic Horseshoe, which is likely a fossil group at $z = 0.44$. Fossil groups, as remnants of early galaxy mergers, may follow distinct evolutionary pathways compared to local galaxies, potentially explaining the high BH mass.

Nonetheless, our analysis found that the stellar mass-to-light ratio and the DM halo of the system are consistent with previous studies of ETGs. The inner DM slope, γ_{DM}, remained consistent with the NFW profile ($\gamma_{DM} = 1$) across most models. Even when incorporating a mass-concentration relation, which yielded slightly steeper DM slopes ($\gamma_{DM} > 1$), the improvements in fit quality were marginal. This support the conclusion that the DM halo of the Cosmic Horseshoe is well-described by an NFW-like profile.

Radial arcs like the one studied here are expected to become increasingly common. The Euclid mission is expected to discover hundreds of thousands of lenses over the next five years (Collett 2015), while the Extremely Large Telescope (ELT) will revolutionize our ability to conduct detailed dynamical studies. The combination of lensing and dynamics will soon provide an unprecedented sample of galaxies, offering exciting insights into stellar populations, DM halos, and SMBHs across a broader redshift range than ever before. This new era of discovery promises to deepen our understanding of galaxy evolution and the interplay between baryonic and DM components.

SOFTWARE CITATIONS

This work uses the following software packages:

Astroalign (Beroiz et al. 2020)
Astropy (Astropy Collaboration et al. 2018)
CMasher (van der Velden 2020)
dynesty (Speagle 2020)
JAM (Cappellari 2008, 2020)
Jupyter (Kluyver et al. 2016)
Matplotlib (Hunter 2007)
MgeFit (Cappellari 2002)
MPDAF (Piqueras et al. 2017)
Numba (Lam et al. 2015)
NumPy (Harris et al. 2020)
pPXF (Cappellari & Emsellem 2004)
PSFr (Birrer et al. 2022)
PyAutoLens (Nightingale et al. 2018)
SciPy (Virtanen et al. 2020)
Shapely (Gillies et al. 2024)
VorBin (Cappellari & Copin 2003)

DATA AVAILABILITY

The HST imaging data is publicly available at the Hubble Legacy Archive (https://hla.stsci.edu/) under the programs 11602 (PI: Sahar Allam) and 12266 (PI: Anna Quider). The MUSE data is available at ESO Science Archive Facility (http://archive.eso.org/scienceportal/) under the program-ID 094.B-0771 (PI: Bethan James). Data relative to Fig. 10 are available at https://github.com/remco-space/Black-Hole-Mass-compilation. Additional data generated in this research will be available upon request to the corresponding author.

REFERENCES

Arenou F., et al., 2018, A&A, 616, A17
Astropy Collaboration et al., 2018, AJ, 156, 123
Barnabè M., et al., 2012, MNRAS, 423, 1073
Beifiori A., Sarzi M., Corsini E. M., Dalla Bontà E., Pizzella A., Coccato L., Bertola F., 2009, ApJ, 692, 856
Belokurov V., et al., 2007, ApJ, 671, L9
Bentz M. C., Katz S., 2015, PASP, 127, 67
Beroiz M., Cabral J., Sanchez B., 2020, Astronomy and Computing, 32, 100384
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition
Birrer S., et al., 2019, MNRAS, 484, 4726
Birrer S., Bhamre V., Nierenberg A., Yang L., Van de Vyvere L., 2022, PSFr: Point Spread Function reconstruction, Astrophysics Source Code Library, record ascl:2210.005
Blandford R. D., McKee C. F., 1982, ApJ, 255, 419
Bogdán Á., Lovisari L., Volonteri M., Dubois Y., 2018, ApJ, 852, 131
Cappellari M., 2002, MNRAS, 333, 400
Cappellari M., 2008, MNRAS, 390, 71
Cappellari M., 2020, MNRAS, 494, 4819
Cappellari M., Copin Y., 2003, MNRAS, 342, 345
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cappellari M., et al., 2013, MNRAS, 432, 1709
Collett T. E., 2015, ApJ, 811, 20
Collett T. E., et al., 2018, Science, 360, 1342
Conroy C., van Dokkum P. G., 2012, ApJ, 760, 71
Del Popolo A., Le Delliou M., 2022, arXiv e-prints, p. arXiv:2209.14151
Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, MNRAS, 437, 415
Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, MNRAS, 433, 3297
Dullo B. T., 2019, ApJ, 886, 80
Dye S., Evans N. W., Belokurov V., Warren S. J., Hewett P., 2008, MNRAS, 388, 384
Emsellem E., Monnet G., Bacon R., 1994, A&A, 285, 723
Emsellem E., et al., 2007, MNRAS, 379, 401
Gebhardt K., et al., 2000, ApJ, 543, L5
Gerhard O. E., 1993, MNRAS, 265, 213
Gerhard O., Kronawitter A., Saglia R. P., Bender R., 2001, AJ, 121, 1936
Gillies S., van der Wel C., Van den Bossche J., Taves M. W., Arnott J., Ward B. C., et al., 2024, Shapely, doi:10.5281/zenodo.13345370, https://doi.org/10.5281/zenodo.13345370
Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16
Gorenstein M. V., Falco E. E., Shapiro I. I., 1988, ApJ, 327, 693
Gu M., Greene J. E., Newman A. B., Kreisch C., Quenneville M. E., Ma C.-P., Blakeslee J. P., 2022, ApJ, 932, 103
Gültekin K., et al., 2009, ApJ, 698, 198
Harris C. R., et al., 2020, Nature, 585, 357
Hlavacek-Larrondo J., Fabian A. C., Edge A. C., Hogan M. T., 2012, MNRAS, 424, 224
Hopkins P. F., Somerville R. S., Hernquist L., Cox T. J., Robertson B., Li Y., 2006, ApJ, 652, 864
Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
Ishibashi W., Fabian A. C., 2012, MNRAS, 427, 2998
Jackson R. A., et al., 2024, MNRAS, 528, 1655
James B., Auger M., Pettini M., Stark D. P., Belokurov V., Carniani S., 2018, in American Astronomical Society Meeting Abstracts #231. p. 210.03
Kluyver T., et al., 2016, in Loizides F., Schmidt B., eds, Positioning and Power in Academic Publishing: Players, Agents and Agendas. pp 87 – 90
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kravtsov A. V., 2013, ApJ, 764, L31
Lam S. K., Pitrou A., Seibert S., 2015, in Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC. LLVM '15. Association for Computing Machinery, New York, NY, USA, doi:10.1145/2833157.2833162, https://doi.org/10.1145/2833157.2833162
Laor A., 2001, ApJ, 553, 677
Liddle A. R., Mukherjee P., Parkinson D., Wang Y., 2006, Phys. Rev. D, 74, 123506
Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214
Magorrian J., et al., 1998, AJ, 115, 2285
Marasco A., Cresci G., Posti L., Fraternali F., Mannucci F., Marconi A., Belfiore F., Fall S. M., 2021, MNRAS, 507, 4274
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Maresca J., Dye S., Li N., 2021, MNRAS, 503, 2229

McConnell N. J., Ma C.-P., Gebhardt K., Wright S. A., Murphy J. D., Lauer T. R., Graham J. R., Richstone D. O., 2011, Nature, 480, 215

McConnell N. J., Ma C.-P., Murphy J. D., Gebhardt K., Lauer T. R., Graham J. R., Wright S. A., Richstone D. O., 2012, ApJ, 756, 179

McLure R. J., Dunlop J. S., 2002, MNRAS, 331, 795

Mehrgan K., Thomas J., Saglia R., Mazzalay X., Erwin P., Bender R., Kluge M., Fabricius M., 2019, ApJ, 887, 195

Melo-Carneiro C. R., Furlanetto C., Chies-Santos A. L., 2023, MNRAS, 520, 1613

Meneghetti M., Bartelmann M., Dahle H., Limousin M., 2013, Space Sci. Rev., 177, 31

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Newman A. B., Ellis R. S., Treu T., 2015, ApJ, 814, 26

Nightingale J. W., Dye S., Massey R. J., 2018, MNRAS, 478, 4738

Nightingale J., et al., 2021, The Journal of Open Source Software, 6, 2825

Nightingale J. W., et al., 2023, MNRAS, 521, 3298

Petit Q., Ducourant C., Slezak E., Sluse D., Delchambre L., 2023, A&A, 669, A132

Piqueras L., Conseil S., Shepherd M., Bacon R., Leclercq F., Richard J., 2017, arXiv e-prints, p. arXiv:1710.03554

Planck Collaboration et al., 2016, A&A, 594, A13

Ponman T. J., Allan D. J., Jones L. R., Merrifield M., McHardy I. M., Lehto H. J., Luppino G. A., 1994, Nature, 369, 462

Powell M. C., et al., 2022, ApJ, 938, 77

Ricciardelli E., Vazdekis A., Cenarro A. J., Falcón-Barroso J., 2012, MNRAS, 424, 172

Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses, doi:10.1007/978-3-662-03758-4.

Schuldt S., Chirivì G., Suyu S. H., Yıldırım A., Sonnenfeld A., Halkola A., Lewis G. F., 2019, A&A, 631, A40

Shen Y., 2013, Bulletin of the Astronomical Society of India, 41, 61

Sibson R., 1981, in Barnett V., ed., , Interpreting Multivariate Data. John Wiley & Sons, New York, pp 21–36

Silk J., 2013, ApJ, 772, 112

Skilling J., 2006, Bayesian Analysis, 1, 833

Smith R. J., Lucey J. R., Edge A. C., 2017a, MNRAS, 467, 836

Smith R. J., Lucey J. R., Edge A. C., 2017b, MNRAS, 471, 383

Sonnenfeld A., Treu T., Marshall P. J., Suyu S. H., Gavazzi R., Auger M. W., Nipoti C., 2015, ApJ, 800, 94

Speagle J. S., 2020, MNRAS, 493, 3132

Spiniello C., Koopmans L. V. E., Trager S. C., Czoske O., Treu T., 2011, MNRAS, 417, 3000

Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, MNRAS, 371, 983

Tessore N., Metcalf R. B., 2015, A&A, 580, A79

Thomas J., Saglia R. P., Bender R., Erwin P., Fabricius M., 2014, ApJ, 782, 39

Thomas J., Ma C.-P., McConnell N. J., Greene J. E., Blakeslee J. P., Janish R., 2016, Nature, 532, 340

Trick W. H., van de Ven G., Dutton A. A., 2016, MNRAS, 463, 3151

Vazdekis A., Ricciardelli E., Cenarro A. J., Rivero-González J. G., Díaz-García L. A., Falcón-Barroso J., 2012, MNRAS, 424, 157

Verro K., et al., 2022, A&A, 661, A50

Virtanen P., et al., 2020, Nature Methods, 17, 261–272

Wang H., et al., 2022, A&A, 668, A162

Warren S. J., Dye S., 2003, ApJ, 590, 673

Weilbacher P. M., Streicher O., Palsa R., 2016, MUSE-DRP: MUSE Data Reduction Pipeline, Astrophysics Source Code Library, record ascl:1610.004

Willmer C. N. A., 2018, ApJS, 236, 47

Wu X., Gerhard O., Naab T., Oser L., Martinez-Valpuesta I., Hilz M., Churazov E., Lyskova N., 2014, MNRAS, 438, 2701

Wu X.-B., et al., 2015, Nature, 518, 512

Wyithe J. S. B., Turner E. L., Spergel D. N., 2001, ApJ, 555, 504

van den Bosch R. C. E., 2016, ApJ, 831, 134

van der Velden E., 2020, The Journal of Open Source Software, 5, 2004

APPENDIX A: CONJUGATED REGIONS

Instead of selecting pairs of conjugated points that are expected to correspond to roughly the same location in the source plane, we opted for choosing pairs of conjugated regions that are expected to overlap at the source plane. This approach is particularly suitable when identifying conjugated points is unclear, such as in the case of faint counter-images. The method is implemented as follows:

First, we select two regions in the image plane that are expected to overlap in the source plane after being delensed. For a given lens macro model, we ray-trace the pixels from the image plane to the source plane. Once mapped to the source plane, we generate convex hull polygons for the individual regions, ensuring all image pixels from the corresponding regions are contained within the delensed regions in the source plane.

We then assess whether the two source plane regions intersect. If they do, the lens macro model is accepted. Otherwise, the log-likelihood is penalised by a factor of $10^8 d_p$, where d_p represents the minimum distance between the two convex hull polygons.

In Fig. A1, we illustrate the method. The left panel shows a pair of conjugated regions in the image plane, which are presumably part of the lensed source. The central panel displays the corresponding regions in the source plane after being delensed. In this scenario, since the regions do not overlap in the source plane, the lens macro model will be penalised. The right panel shows the case where the regions overlap in the source plane after being delensed. In this scenario, the lens macro model is accepted.

We used the `ConvexHull` routine from `scipy.spatial` to create the convex hull polygons in the source plane, and the `shapely` package to calculate the intersection and distance between the source plane polygons.

APPENDIX B: ELLIPTICAL POWER-LAW (EPL) MASS MODEL

The EPL density profile (Tessore & Metcalf 2015) is widely employed in SGL studies to characterise the total mass distribution of the lens. The convergence for this profile is expressed as:

$$\kappa(\xi) = \frac{\left(3 - \gamma^{\text{lens}}\right)}{1 + q^{\text{lens}}} \left(\frac{\theta_{\text{Ein}}^{\text{lens}}}{\xi}\right)^{\gamma^{\text{lens}} - 1}, \tag{B1}$$

where q^{lens} the axis ratio (minor-to-major axis), and ξ is the elliptical coordinate given by $\xi = \sqrt{x^2 + (y/q^{\text{lens}})^2}$. The parameter $\theta_{\text{Ein}}^{\text{lens}}$ is Einstein radius in units of arcsec, and γ^{lens} is the mass density slope, which reduces to a singular isothermal ellipsoid (SIE) for $\gamma^{\text{lens}} = 2$.

Additionally, the mass position angle, ϕ^{lens}, measured counterclockwise from the positive x-axis, can be incorporated by introducing the elliptical components:

$$\epsilon_1 = \frac{1 - q^{\text{lens}}}{1 + q^{\text{lens}}} \sin 2\phi^{\text{lens}}, \qquad \epsilon_2 = \frac{1 - q^{\text{lens}}}{1 + q^{\text{lens}}} \cos 2\phi^{\text{lens}}. \tag{B2}$$

It is essential to distinguish between the Einstein radius $\theta_{\text{Ein}}^{\text{lens}}$ used in this equation and the *effective* Einstein radius, as defined in Meneghetti et al. (2013). The *effective* Einstein radius corresponds to the radius of a circle with the same area as the region enclosed by the tangential critical curve. The Einstein radius reported in Section 3.1 refers to the effective definition.

We modelled the lensed source s2 in the Cosmic Horseshoe using the EPL mass profile to describe the total mass distribution of
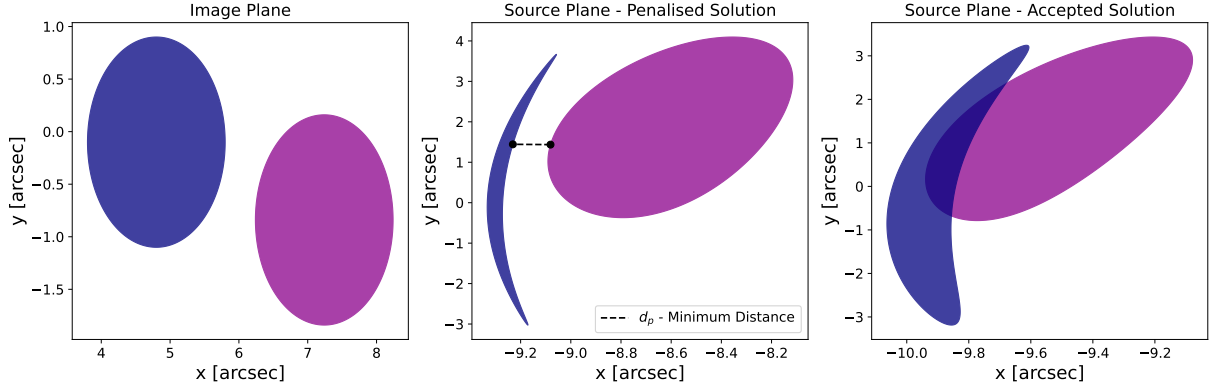
Figure A1. Schematic representation of the conjugated region method. The left panel shows the pair of conjugated regions in the image plane. The central panel illustrates a scenario where the lens macro model is rejected because the delensed regions do not overlap in the source plane. In this case, the likelihood is penalised by a factor proportional to the minimum distance between the polygons, represented by the dashed line. The right panel depicts the case where the regions overlap in the source plane after inversion, leading to the acceptance of the lens macro model.

Table B1. Inferred median and 1σ credible intervals for the parameters of the EPL mass model.

Parameter	Posterior (median with 1σ uncertainties)
ϵ_1	$-0.097^{+0.001}_{-0.001}$
ϵ_2	$-0.013^{+0.01}_{-0.001}$
$\theta^{\text{lens}}_{\text{Ein}}[''] $	$5.002^{+0.009}_{-0.005}$
γ^{lens}	$1.82^{+0.02}_{-0.01}$
ϵ^{sh}_1	$-0.026^{+0.001}_{-0.001}$
ϵ^{sh}_2	$-0.008^{+0.001}_{-0.001}$

the main deflector. We also added an external shear contribution. For the source reconstruction, we utilised a `KMeans` mesh grid with `VoronoiNN` pixelisation and a `AdaptiveBrightnessSplit` regularisation, which adapts the smoothing based on the source's surface brightness. Aditionally, during the lens modelling, we only included pixels within the mask encompassing the arc, as illustrated in the middle panel of Fig. 3.

The highest-likelihood image model and source reconstruction are presented in Fig. 3. The median values and associated 1σ uncertainties of the parameter's one-dimensional marginalised posterior of the EPL model are summarised in Table B1.

APPENDIX C: PRIORS

In Table C1 we describe the parameters and the priors applied in the models presented in this work, and discussed in Section 4. We adopted the following notation: $U[a, b]$ for a uniform prior between the lower value a and the upper value b; $N[a, b]$ for a normal Gaussian prior with mean a and dispersion b; and $\log_{10} U[a, b]$ for a log-uniform prior between the lower value a and the upper value b.

APPENDIX D: ALTERNATIVE MASS MODELS RESULTS

The median and uncertainties of the parameters in the pertubation models are summarised in Table D1. Units and priors can be checked in Table C1.

In Table D2 we present the results of the alternative model **M9**.
In Table D3 we present the results of the alternative model **M10**.
In Table D4 we present the results of the alternative models **M11** — **M13**, where we used the mass-concentration relation from Ludlow et al. (2016) to set the DM scale radius.
In Table D5 we present the results of the alternative models **M14** and **M15** in which we did not include the SMBH.

This paper has been typeset from a TEX/LATEX file prepared by the author.

Table C1. Parameters and priors used in this work. From left to right, the columns are: parameter, prior, parameter description, and physical unit. [†] Minimum value is determined by the minimum axial ratio allowed by Eq. 14. [‡] Relative to the mass-to-light ratio profile described by Eq. 17.

Parameter	Prior	Description	Physical Unit
i	$U[49.55, 90]^{\dagger}$	inclination along the line-of-sight	degree
$\beta_{\rm star}$	$U[-0.5, 0.5]$	stellar anisotropy	-
$\Upsilon_{\star}$	$U[0.1, 10]$	stellar mass-to-light ratio	$M_{\odot}/L_{\odot}$
$\log_{10} \rho_s$	$U[-6, 0]$	dark matter characteristic density	$\frac{\rho_s}{M_{\odot}{\rm pc}^{-3}}$
$M_{200}^{\rm DM}$	$\log_{10} U[10^{10}, 10^{15}]$	mass at 200 times the critical density of the Universe	$M_{\odot}$
r_s	$U[10, 30]$	dark matter scale radius	arcsec
$q_{\rm DM}$	$U[0.65, 1]$	dark matter axial ratio	-
$\gamma_{\rm DM}$	$U[0, 2]$	dark matter density slope	-
$\log_{10} M_{\rm BH}$	$\log_{10} U[8, 12]$	mass of the SMBH	$\frac{M_{\rm BH}}{M_{\odot}}$
Υ_0	$U[0.1, 10]$	central stellar mass-to-light ratio[‡]	$M_{\odot}/L_{\odot}$
υ_0	$U[0, 10]$	ratio between the central and outermost stellar mass-to-light ratio[‡]	-
δ	$U[0.1, 1]$	smoothness of the stellar mass-to-light ratio[‡]	arcsec^{-1}
$\epsilon_1^{\rm sh}$	$U[-0.2, 0.2]$	elliptical shear component	-
$\epsilon_2^{\rm sh}$	$U[-0.2, 0.2]$	elliptical shear component	-

Table D1. Posterior median and 1σ uncertainties for the parameters in the perturbation models.

Parameter	M2	M3	M4	M5	M6	M7	M8
i	66^{+14}_{-10}	61^{+11}_{-7}	67^{+15}_{-12}	59^{+5}_{-5}	64^{+17}_{-11}	66^{+14}_{-11}	69^{+13}_{-12}
$\beta^0_{\rm star}$	$-0.14^{+0.20}_{-0.20}$	$0.04^{+0.34}_{-0.34}$	$0.09^{+0.12}_{-0.15}$	$-0.14^{+0.16}_{-0.14}$	$0.03^{+0.05}_{-0.05}$	$-0.01^{+0.10}_{-0.16}$	$-0.17^{+0.14}_{-0.16}$
$\beta^1_{\rm star}$	-	$0.35^{+0.10}_{-0.18}$	-	-	-	-	-
$\beta^2_{\rm star}$	-	$-0.01^{+0.08}_{-0.08}$	-	-	-	-	-
$\Upsilon^0_\star$	$3.24^{+0.28}_{-0.29}$	$3.54^{+0.27}_{-0.35}$	$6.94^{+1.97}_{-1.70}$	$3.35^{+0.30}_{-0.35}$	$3.24^{+0.29}_{-0.33}$	$3.25^{+0.35}_{-0.33}$	$3.65^{+0.33}_{-0.28}$
$\Upsilon^1_\star$	-	-	$3.09^{+0.30}_{-0.26}$	-	-	-	-
$\Upsilon^2_\star$	-	-	$2.52^{+0.49}_{-0.65}$	-	-	-	-
$\log_{10}\rho_s$	$-2.43^{+0.05}_{-0.04}$	$-2.38^{+0.01}_{-0.01}$	$-2.36^{+0.01}_{-0.01}$	$-2.32^{+0.04}_{-0.07}$	$-2.38^{+0.01}_{-0.01}$	$-2.38^{+0.01}_{-0.01}$	$-2.39^{+0.01}_{-0.01}$
r_s	-	-	-	$19.16^{+2.30}_{-0.97}$	-	-	-
$q_{\rm DM}$	$0.98^{+0.01}_{-0.01}$	$0.96^{+0.01}_{-0.01}$	$0.96^{+0.02}_{-0.01}$	$0.97^{+0.02}_{-0.02}$	$0.98^{+0.01}_{-0.02}$	$0.97^{+0.01}_{-0.01}$	$0.93^{+0.05}_{-0.06}$
$\gamma_{\rm DM}$	$1.06^{+0.05}_{-0.07}$	-	-	-	-	-	-
$\log_{10} M_{\rm BH}$	$10.57^{+0.07}_{-0.09}$	$10.45^{+0.11}_{-0.14}$	$10.53^{+0.10}_{-0.11}$	$10.56^{+0.08}_{-0.08}$	$10.55^{+0.08}_{-0.09}$	$10.55^{+0.08}_{-0.08}$	$10.51^{+0.07}_{-0.09}$
$\epsilon^{\rm sh}_1$	$0.02^{+0.01}_{-0.01}$	$0.03^{+0.01}_{-0.01}$	$0.03^{+0.01}_{-0.01}$	$0.03^{+0.01}_{-0.01}$	$0.02^{+0.01}_{-0.01}$	$0.03^{+0.01}_{-0.01}$	$0.02^{+0.02}_{-0.01}$
$\epsilon^{\rm sh}_2$	$0.05^{+0.01}_{-0.01}$	$0.06^{+0.01}_{-0.01}$	$0.05^{+0.01}_{-0.01}$	$0.06^{+0.01}_{-0.01}$	$0.05^{+0.01}_{-0.01}$	$0.05^{+0.02}_{-0.01}$	$0.05^{+0.01}_{-0.01}$
$M_{\rm Ein}\left[\frac{M_\odot}{10^{12}}\right]$	$5.45^{+0.03}_{-0.03}$	$5.45^{+0.02}_{-0.02}$	$5.45^{+0.02}_{-0.02}$	$5.45^{+0.02}_{-0.02}$	$5.46^{+0.02}_{-0.03}$	$5.45^{+0.02}_{-0.02}$	-

Table D2. Inferred median and 1σ credible intervals for the parameters of model **M9**

Parameter	Posterior (median with 1σ uncertainties)
i	71^{+10}_{-8}
$\beta_{\rm star}^0$	$-0.21^{+0.17}_{-0.13}$
$\beta_{\rm star}^1$	$-0.05^{+0.41}_{-0.15}$
$\beta_{\rm star}^2$	$-0.02^{+0.16}_{-0.17}$
$\Upsilon_\star^0$	$5.17^{+1.19}_{-0.84}$
$\Upsilon_\star^1$	$3.21^{+0.79}_{-0.48}$
$\Upsilon_\star^2$	$1.66^{+1.87}_{-0.83}$
$\log_{10}\rho_s$	$-2.41^{+0.04}_{-0.05}$
$q_{\rm DM}$	$0.97^{+0.03}_{-0.02}$
$\gamma_{\rm DM}$	$1.08^{+0.06}_{-0.07}$
$\log_{10} M_{\rm BH}$	$10.50^{+0.10}_{-0.32}$
$\epsilon_1^{\rm sh}$	$0.03^{+0.01}_{-0.01}$
$\epsilon_2^{\rm sh}$	$0.06^{+0.01}_{-0.01}$
$M_{\rm Ein}\left[\frac{M_\odot}{10^{12}}\right]$	$5.45^{+0.02}_{-0.03}$

Table D3. Inferred median and 1σ credible intervals for the parameters of model **M10**

Parameter	Posterior (median with 1σ uncertainties)
i	75^{+10}_{-8}
$\beta_{\rm star}^0$	$-0.07^{+0.21}_{-0.62}$
$\beta_{\rm star}^1$	$-0.51^{+0.87}_{-0.40}$
$\beta_{\rm star}^2$	$-0.34^{+0.36}_{-0.39}$
$\beta_{\rm star}^3$	$0.12^{+0.18}_{-0.32}$
$\beta_{\rm star}^4$	$0.37^{+0.16}_{-0.42}$
$\beta_{\rm star}^5$	$0.05^{+0.21}_{-0.27}$
$\beta_{\rm star}^6$	$-0.43^{+0.13}_{-0.17}$
$\beta_{\rm star}^7$	$0.13^{+0.09}_{-0.05}$
Υ_0	$4.38^{+0.57}_{-1.31}$
υ_0	$0.53^{+0.08}_{-0.08}$
δ	$7.77^{+1.72}_{-7.20}$
$\log_{10}\rho_s$	$-2.38^{+0.01}_{-0.01}$
$\log_{10} M_{\rm BH}$	$10.55^{+0.10}_{-0.07}$
$\epsilon_1^{\rm sh}$	$0.00^{+0.01}_{-0.01}$
$\epsilon_2^{\rm sh}$	$0.05^{+0.01}_{-0.01}$
$M_{\rm Ein}\left[\frac{M_\odot}{10^{12}}\right]$	$5.45^{+0.02}_{-0.02}$

Table D4. Inferred median and 1σ credible intervals for the parameters of models **M11 — M13**. Note that the DM scale radius is not a free parameter in these models, but setted using the mass-concentration relation from Ludlow et al. (2016).

Parameter	M11	M12	M13
i	75^{+8}_{-10}	76^{+9}_{-12}	65^{+20}_{-11}
$\beta_{\rm star}^0$	$-0.01^{+0.24}_{-0.24}$	$-0.02^{+0.24}_{-0.28}$	$0.08^{+0.25}_{-0.15}$
$\beta_{\rm star}^1$	$0.42^{+0.06}_{-0.11}$	$0.34^{+0.09}_{-0.10}$	$-0.39^{+0.22}_{-0.07}$
$\beta_{\rm star}^2$	$-0.29^{+0.19}_{-0.14}$	$-0.10^{+0.24}_{-0.23}$	$-0.26^{+0.16}_{-0.11}$
$\Upsilon_\star^0$	$5.39^{+1.81}_{-1.33}$	$5.79^{+2.40}_{-1.64}$	$4.51^{+1.24}_{-0.89}$
$\Upsilon_\star^1$	$3.63^{+0.32}_{-0.35}$	$3.14^{+0.16}_{-0.22}$	$3.16^{+0.34}_{-0.20}$
$\Upsilon_\star^2$	$2.75^{+0.61}_{-1.04}$	$1.55^{+0.70}_{-0.77}$	$2.69^{+0.48}_{-0.62}$
$M_{200}^{\rm DM}\left[\frac{M_\odot}{10^{13}}\right]$	$8.36^{+1.48}_{-1.22}$	$11.1^{+4.36}_{-2.84}$	$9.52^{+1.28}_{-7.92}$
$q_{\rm DM}$	$0.99^{+0.01}_{-0.01}$	$0.99^{+0.01}_{-0.01}$	$0.99^{+0.01}_{-0.01}$
r_s	$27.96^{+2.10}_{-1.90}$	$44.89^{+7.14}_{-5.51}$	$20.97^{+1.21}_{-0.80}$
$\gamma_{\rm DM}$	$1.29^{+0.05}_{-0.04}$	$1.34^{+0.07}_{-0.07}$	$1.12^{+0.03}_{-0.04}$
$\log_{10} M_{\rm BH}$	$10.15^{+0.17}_{-0.30}$	$10.33^{+0.07}_{-0.13}$	$10.59^{+0.04}_{-0.10}$
$\epsilon_1^{\rm sh}$	$0.03^{+0.01}_{-0.01}$	$0.03^{+0.01}_{-0.01}$	$0.01^{+0.01}_{-0.01}$
$\epsilon_2^{\rm sh}$	$0.06^{+0.01}_{-0.01}$	$0.06^{+0.01}_{-0.01}$	$0.06^{+0.01}_{-0.01}$
$M_{\rm Ein}\left[\frac{M_\odot}{10^{12}}\right]$	$5.44^{+0.02}_{-0.02}$	$5.43^{+0.02}_{-0.02}$	$5.43^{+0.02}_{-0.02}$

Table D5. Inferred median and 1σ credible intervals for the parameters of models without the SMBH.

Parameter	M14	M15
i	79^{+7}_{-9}	79^{+7}_{-10}
$\beta_{\rm star}$	$0.28^{+0.04}_{-0.04}$	$0.42^{+0.05}_{-0.08}$
$\Upsilon_\star^0$	$4.14^{+0.08}_{-0.07}$	$7.57^{+1.71}_{-2.26}$
$\Upsilon_\star^1$	-	$3.90^{+0.19}_{-0.13}$
$\Upsilon_\star^2$	-	$1.29^{+0.90}_{-0.70}$
$\log_{10}\rho_s$	$-2.40^{+0.01}_{-0.01}$	$-2.36^{+0.01}_{-0.02}$
$q_{\rm DM}$	$0.99^{+0.01}_{-0.01}$	$0.98^{+0.01}_{-0.01}$
$\epsilon_1^{\rm sh}$	$0.02^{+0.01}_{-0.01}$	$0.02^{+0.01}_{-0.01}$
$\epsilon_2^{\rm sh}$	$0.05^{+0.01}_{-0.01}$	$0.05^{+0.01}_{-0.01}$
$M_{\rm Ein}\left[\frac{M_\odot}{10^{12}}\right]$	$5.45^{+0.02}_{-0.03}$	$5.45^{+0.03}_{-0.03}$